

# (11) **EP 1 790 660 A2**

(12)

# **EUROPEAN PATENT APPLICATION**

(43) Date of publication: 30.05.2007 Bulletin 2007/22

(51) Int Cl.: C07K 14/22 (2006.01)

(21) Application number: 06076718.3

(22) Date of filing: 28.02.2001

(84) Designated Contracting States:

AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU

MC NL PT SE TR

(30) Priority: **28.02.2000 GB 0004695 13.11.2000 GB 0027675** 

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC: 01914109.2 / 1 259 627

(71) Applicant: Novartis Vaccines and Diagnostics S.r.I.53100 Siena SI (IT)

(72) Inventors:

- Arico, Maria B.
   53100 Siena (IT)
- Comanducci, Maurizio 53100 Siena (IT)
- Galeotti, Cesira
   53100 Siena (IT)

- Masignani, Vega
   53100 Siena (IT)
- Guiliani, Marzia Monica 53100 Siena (IT)
- Pizza, Mariagrazia 53100 Siena (IT)
- (74) Representative: Marshall, Cameron John et al Carpmaels & Ransford
   43-45 Bloomsbury Square London WC1A 2RA (GB)

## Remarks:

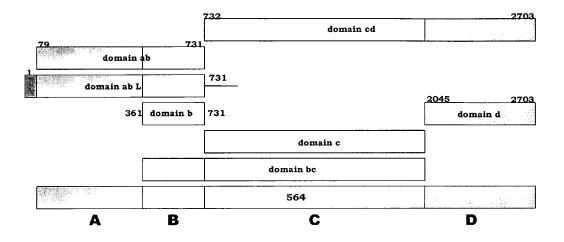
- •This application was filed on 13 09 2006 as a divisional application to the application mentioned under INID code 62.
- •The sequence listing, which is published as annex to the application documents, was filed after the date of filing. The applicant has declared that it does not include matter which goes beyond the content of the application as filed.

## (54) Heterologous expression of neisserial proteins

(57) Alternative approaches to the heterologous expression of the proteins of *Neisseria meningitidis* and *Neisseria gonorrhoeae*. These approaches typically af-

fect the level of expression, the ease of purification, the cellular localisation, and/or the immunological properties of the expressed protein.

## FIGURE 8



#### Description

#### **TECHNICAL FIELD**

[0001] This invention is in the field of protein expression. In particular, it relates to the heterologous expression of proteins from *Neisseria* (e.g. *N.gonorrhoeae* or, preferably, *N.meningitidis*).

#### **BACKGROUND ART**

[0002] International patent applications WO99/24578, WO99/36544, WO99/57280 and WO00/22430 disclose proteins from Neisseria meningitidis and Neisseria gonorrhoeae. These proteins are typically described as being expressed in E.coli (i. e. heterologous expression) as either N-terminal GST-fusions or C-terminal His-tag fusions, although other expression systems, including expression in native Neisseria, are also disclosed.

**[0003]** It is an object of the present invention to provide alternative and improved approaches for the heterologous expression of these proteins. These approaches will typically affect the level of expression, the ease of purification, the cellular localisation of expression, and/or the immunological properties of the expressed protein.

#### **DISCLOSURE OF THE INVENTION**

#### Nomenclature herein

20

25

30

35

40

45

50

[0004] The 2166 protein sequences disclosed in WO99/24578, WO99/36544 and WO99/57280 are referred to herein by the following SEQ# numbers:

Application	Protein sequences	SEQ# herein
WO99/24578	Even SEQ IDs 2-892	SEQ#s 1-446
WO99/36544	SEQ#s 447-491	
	Even SEQ IDs 2-3020	SEQ#s 492-2001
WO99/57280	Even SEQ IDs 3040-3114	SEQ#s 2002-2039
	SEQ IDs 3115-3241	SEQ#s 2040-2166

**[0005]** In addition to this SEQ# numbering, the naming conventions used in WO99/24578, WO99/36544 and WO99/57280 are also used (*e.g.* 'ORF4', 'ORF40', 'ORF40-1' *etc.* as used in WO99/24578 and WO99/36544; 'm919', 'g919' and 'a919' *etc.* as used in WO99/57280).

**[0006]** The 2160 proteins NMB0001 to NMB2160 from Tettelin et al. [Science (2000) 287:1809-1815] are referred to herein as SEQ#s 2167-4326 [see also WO00/66791].

[0007] The term 'protein of the invention' as used herein refers to a protein comprising:

- (a) one of sequences SEQ#s 1-4326; or
- (b) a sequence having sequence identity to one of SEQ#s 1-4326; or
- (c) a fragment of one of SEQ#s 1-4326.

[0008] The degree of 'sequence identity' referred to in (b) is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more). This includes mutants and allelic variants [e.g. see WO00/66741]. Identity is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), using an affine gap search with parameters gap open penalty=12 and gap extension penalty=1. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence.

[0009] The 'fragment' referred to in (c) should comprise at least n consecutive amino acids from one of SEQ#s 1-4326 and, depending on the particular sequence, n is 7 or more (eg. 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 or more). Preferably the fragment comprises an epitope from one of SEQ#s 1-4326. Preferred fragments are those disclosed in WO00/71574 and WO01/04316.

[0010] Preferred proteins of the invention are found in *N.meningitidis* serogroup B.

**[0011]** Preferred proteins for use according to the invention are those of serogroup B *N.meningitidis* strain 2996 or strain 394/98 (a New Zealand strain). Unless otherwise stated, proteins mentioned herein are from *N.meningitidis* strain 2996. It will be appreciated, however, that the invention is not in general limited by strain. References to a particular

55

protein (e.g. '287', '919' etc.) may be taken to include that protein from any strain.

#### Non-fusion expression

10

25

30

35

40

50

55

[0012] In a first approach to heterologous expression, no fusion partner is used, and the native leader peptide (if present) is used. This will typically prevent any 'interference' from fusion partners and may alter cellular localisation and/or post-translational modification and/or folding in the heterologous host.

[0013] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) no fusion partner is used, and (b) the protein's native leader peptide (if present) is used.

[0014] The method will typically involve the step of preparing an vector for expressing a protein of the invention, such that the first expressed amino acid is the first amino acid (methionine) of said protein, and last expressed amino acid is the last amino acid of said protein (*i.e.* the codon preceding the native STOP codon).

[0015] This approach is preferably used for the expression of the following proteins using the native leader peptide: 111, 149, 206, 225-1, 235, 247-1, 274, 283, 286, 292, 401, 406, 502-1, 503, 519-1, 525-1, 552, 556, 557, 570, 576-1, 580, 583, 664, 759, 907, 913, 920-1, 936-1, 953, 961, 983, 989, Orf4, Orf7-1, Orf9-1, Orf23, Orf25, Orf37, Orf38, Orf40, Orf40.1, Orf40.2, Orf72-1, Orf76-1, Orf85-2, Orf91, Orf97-1, Orf119, Orf143.1, NMB0109 and NMB2050. The suffix 'L' used herein in the name of a protein indicates expression in this manner using the native leader peptide.

**[0016]** Proteins which are preferably expressed using this approach using no fusion partner and which have no native leader peptide include: 008, 105, 117-1, 121-1, 122-1, 128-1, 148, 216, 243, 308, 593, 652, 726, 926, 982, Orf83-1 and Orf143-1.

[0017] Advantageously, it is used for the expression of ORF25 or ORF40, resulting in a protein which induces better anti-bactericidal antibodies than GST- or His-fusions.

[0018] This approach is particularly suited for expressing lipoproteins.

## Leader-peptide substitution

**[0019]** In a second approach to heterologous expression, the native leader peptide of a protein of the invention is replaced by that of a different protein. In addition, it is preferred that no fusion partner is used. Whilst using a protein's own leader peptide in heterologous hosts can often localise the protein to its 'natural' cellular location, in some cases the leader sequence is not efficiently recognised by the heterologous host. In such cases, a leader peptide known to drive protein targeting efficiently can be used instead.

[0020] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) the protein's leader peptide is replaced by the leader peptide from a different protein and, optionally, (b) no fusion partner is used.

[0021] The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; manipulating said nucleic acid to remove nucleotides that encode the protein's leader peptide and to introduce nucleotides that encode a different protein's leader peptide. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector. The expressed protein will consist of the replacement leader peptide at the N-terminus, followed by the protein of the invention minus its leader peptide.

[0022] The leader peptide is preferably from another protein of the invention (e.g. one of SEQ#s 1-4326), but may also be from an *E.coli* protein (*e.g.* the OmpA leader peptide) or an *Erwinia carotovora* protein (*e.g.* the PelB leader peptide), for instance.

**[0023]** A particularly useful replacement leader peptide is that of ORF4. This leader is able to direct lipidation in *E.coli*, improving cellular localisation, and is particularly useful for the expression of proteins 287, 919 and  $\Delta$ G287. The leader peptide and N-terminal domains of 961 are also particularly useful.

**[0024]** Another useful replacement leader peptide is that of *E.coli* OmpA. This leader is able to direct membrane localisation of *E.coli*. It is particularly advantageous for the expression of ORF1, resulting in a protein which induces better anti-bactericidal antibodies than both fusions and protein expressed from its own leader peptide.

**[0025]** Another useful replacement leader peptide is MKKYLFSAA. This can direct secretion into culture medium, and is extremely short and active. The use of this leader peptide is not restricted to the expression of Neisserial proteins - it may be used to direct the expression of any protein (particularly bacterial proteins).

#### Leader-peptide deletion

[0026] In a third approach to heterologous expression, the native leader peptide of a protein of the invention is deleted. In addition, it is preferred that no fusion partner is used.

[0027] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) the protein's leader peptide is deleted and, optionally, (b) no fusion partner is used.

**[0028]** The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; manipulating said nucleic acid to remove nucleotides that encode the protein's leader peptide. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector. The first amino acid of the expressed protein will be that of the mature native protein.

**[0029]** This method can increase the levels of expression. For protein 919, for example, expression levels in *E.coli* are much higher when the leader peptide is deleted. Increased expression may be due to altered localisation in the absence of the leader peptide.

[0030] The method is preferably used for the expression of 919, ORF46, 961, 050-1, 760 and 287.

#### Domain-based expression

10

30

35

40

45

50

[0031] In a fourth approach to heterologous expression, the protein is expressed as domains. This may be used in association with fusion systems (e.g. GST or His-tag fusions).

[0032] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) at least one domain in the protein is deleted and, optionally, (b) no fusion partner is used.

**[0033]** The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; manipulating said nucleic acid to remove at least one domain from within the protein. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector. Where no fusion partners are used, the first amino acid of the expressed protein will be that of a domain of the protein.

[0034] A protein is typically divided into notional domains by aligning it with known sequences in databases and then determining regions of the protein which show different alignment patterns from each other.

**[0035]** The method is preferably used for the expression of protein 287. This protein can be notionally split into three domains, referred to as A B & C (see Figure 5). Domain B aligns strongly with IgA proteases, domain C aligns strongly with transferrin-binding proteins, and domain A shows no strong alignment with database sequences. An alignment of polymorphic forms of 287 is disclosed in WO00/66741.

**[0036]** Once a protein has been divided into domains, these can be (a) expressed singly (b) deleted from with the protein e.g. protein ABCD  $\rightarrow$  ABD, ACD, BCD *etc.* or (c) rearranged e.g. protein ABC  $\rightarrow$  ACB, CAB *etc.* These three strategies can be combined with fusion partners is desired.

[0037] ORF46 has also been notionally split into two domains - a first domain (amino acids 1-433) which is well-conserved between species and serogroups, and a second domain (amino acids 433-608) which is not well-conserved. The second domain is preferably deleted. An alignment of polymorphic forms of ORF46 is disclosed in WO00/66741.

[0038] Protein 564 has also been split into domains (Figure 8), as have protein 961 (Figure 12) and protein 502 (amino acids 28-167 of the MC58 protein).

## Hybrid proteins

[0039] In a fifth approach to heterologous expression, two or more (e.g. 3, 4, 5, 6 or more) proteins of the invention are expressed as a single hybrid protein. It is preferred that no non-Neisserial fusion partner (e.g. GST or poly-His) is used. [0040] This offers two advantages. Firstly, a protein that may be unstable or poorly expressed on its own can be assisted by adding a suitable hybrid partner that overcomes the problem. Secondly, commercial manufacture is simplified - only one expression and purification need be employed in order to produce two separately-useful proteins.

[0041] Thus the invention provides a method for the simultaneous heterologous expression of two or more proteins of the invention, in which said two or more proteins of the invention are fused (i.e. they are translated as a single polypeptide chain).

**[0042]** The method will typically involve the steps of: obtaining a first nucleic acid encoding a first protein of the invention; obtaining a second nucleic acid encoding a second protein of the invention; ligating the first and second nucleic acids. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector.

[0043] Preferably, the constituent proteins in a hybrid protein according to the invention will be from the same strain.

**[0044]** The fused proteins in the hybrid may be joined directly, or may be joined via a linker peptide e.g. via a polyglycine linker (*i.e.*  $G_n$  where n = 3, 4, 5, 6, 7, 8, 9, 10 or more) or via a short peptide sequence which facilitates cloning. It is evidently preferred not to join a  $\Delta G$  protein to the C-terminus of a poly-glycine linker.

[0045] The fused proteins may lack native leader peptides or may include the leader peptide sequence of the N-terminal fusion partner.

[0046] The method is well suited to the expression of proteins orf1, orf4, orf25, orf40, Orf46/46.1, orf83, 233, 287, 292L, 564, 687, 741, 907, 919, 953, 961 and 983.

[0047] The 42 hybrids indicated by 'X' in the following table of form NH<sub>2</sub>-A-B-COOH are preferred:

↓A	$B{ o}$	ORF46.1	287	741	919	953	961	983
ORF	46.1		х	х	Х	Х	Х	Х
	287	Х		х	Х	Х	х	Х
	741	Х	х		Х	Х	Х	Х
	919	Х	х	х		Х	Х	Х
	953	Х	х	х	Х		х	Х
	961	Х	Х	Х	Х	Х		Х
	983	х	Х	х	Х	Х	х	

[0048] Preferred proteins to be expressed as hybrids are thus ORF46.1, 287, 741, 919, 953, 961 and 983. These may be used in their essentially full-length form, or poly-glycine deletions ( $\Delta$ G) forms may be used (e.g.  $\Delta$ G-287,  $\Delta$ GTbp2,  $\Delta$ G741,  $\Delta$ G983 etc.), or truncated forms may be used (e.g.  $\Delta$ 1-287,  $\Delta$ 2-287 etc.), or domain-deleted versions may be used (e.g. 287B, 287C, 287BC, ORF46<sub>1-433</sub>, ORF46<sub>433-608</sub>, ORF46, 961c etc.).

**[0049]** Particularly preferred are: (a) a hybrid protein comprising 919 and 287; (b) a hybrid protein comprising 953 and 287; (c) a hybrid protein comprising 287 and ORF46.1; (d) a hybrid protein comprising ORF1 and ORF46.1; (e) a hybrid protein comprising 919 and ORF46.1; (f) a hybrid protein comprising ORF46.1 and 919; (g) a hybrid protein comprising ORF46.1, 287 and 919; (h) a hybrid protein comprising 919 and 519; and (i) a hybrid protein comprising ORF97 and 225. Further embodiments are shown in Figure 14.

**[0050]** Where 287 is used, it is preferably at the C-terminal end of a hybrid; if it is to be used at the N-terminus, if is preferred to use a  $\Delta G$  form of 287 is used (e.g. as the N-terminus of a hybrid with ORF46.1, 919, 953 or 961).

[0051] Where 287 is used, this is preferably from strain 2996 or from strain 394/98.

[0052] Where 961 is used, this is preferably at the N-terminus. Domain forms of 961 may be used.

**[0053]** Alignments of polymorphic forms of ORF46, 287, 919 and 953 are disclosed in WO00/66741. Any of these polymorphs can be used according to the present invention.

## Temperature

5

10

30

40

50

55

[0054] In a sixth approach to heterologous expression, proteins of the invention are expressed at a low temperature.

**[0055]** Expressed Neisserial proteins (e.g. 919) may be toxic to *E.coli*, which can be avoided by expressing the toxic protein at a temperature at which its toxic activity is not manifested.

**[0056]** Thus the present invention provides a method for the heterologous expression of a protein of the invention, in which expression of a protein of the invention is carried out at a temperature at which a toxic activity of the protein is not manifested.

[0057] A preferred temperature is around 30°C. This is particularly suited to the expression of 919.

## Mutations

**[0058]** As discussed above, expressed Neisserial proteins may be toxic to *E coli*. This toxicity can be avoided by mutating the protein to reduce or eliminate the toxic activity. In particular, mutations to reduce or eliminate toxic enzymatic activity can be used, preferably using site-directed mutagenesis.

[0059] In a seventh approach to heterologous expression, therefore, an expressed protein is mutated to reduce or eliminate toxic activity.

**[0060]** Thus the invention provides a method for the heterologous expression of a protein of the invention, in which protein is mutated to reduce or eliminate toxic activity.

**[0061]** The method is preferably used for the expression of protein 907, 919 or 922. A preferred mutation in 907 is at Glu-117 (e.g. Glu→Gly); preferred mutations in 919 are at Glu-255 (e.g. Glu→Gly) and/or Glu-323 (e.g. Glu→Gly); preferred mutations in 922 are at Glu-164 (e.g. Glu→Gly), Ser-213 (e.g. Ser→Gly) and/or Asn-348 (e.g. Asn→Gly).

#### Alternative vectors

**[0062]** In a eighth approach to heterologous expression, an alternative vector used to express the protein. This may be to improve expression yields, for instance, or to utilise plasmids that are already approved for GMP use.

[0063] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which

an alternative vector is used. The alternative vector is preferably pSM214, with no fusion partners. Leader peptides may or may not be included.

**[0064]** This approach is particularly useful for protein 953. Expression and localisation of 953 with its native leader peptide expressed from pSM214 is much better than from the pET vector.

[0065] pSM214 may also be used with:  $\Delta$ G287,  $\Delta$ 2-287,  $\Delta$ 3-287,  $\Delta$ 4-287, Orf46.1, 961L, 961, 96 1 (MC58), 96 1 c, 96 1 c-L, 919, 953 and  $\Delta$ G287-Orf46.1.

**[0066]** Another suitable vector is pET-24b (Novagen; uses kanamycin resistance), again using no fusion partners. pET-24b is preferred for use with:  $\Delta$ G287K,  $\Delta$ 2-287K,  $\Delta$ 3-287K,  $\Delta$ 4-287K,

[0067] Orf46.1-K, Orf46A-K, 961-K (MC58), 961a-K, 961b-K, 961c-L, 961c-L-K, 961d-K, ΔG287-919-K, ΔG287-Orf46.1-K and ΔG287-961-K.

#### Multimeric form

10

30

35

40

45

50

55

[0068] In a ninth approach to heterologous expression, a protein is expressed or purified such that it adopts a particular multimeric form.

**[0069]** This approach is particularly suited to protein 953. Purification of one particular multimeric form of 953 (the monomeric form) gives a protein with greater bactericidal activity than other forms (the dimeric form).

[0070] Proteins 287 and 919 may be purified in dimeric forms.

[0071] Protein 961 may be purified in a 180kDa oligomeric form (e.g. a tetramer).

## Lipidation

[0072] In a tenth approach to heterologous expression, a protein is expressed as a lipidated protein.

[0073] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which the protein is expressed as a lipidated protein.

[0074] This is particularly useful for the expression of 919, 287, ORF4, 406, 576-1, and ORF25. Polymorphic forms of 919, 287 and ORF4 are disclosed in WO00/66741.

[0075] The method will typically involve the use of an appropriate leader peptide without using an N-terminal fusion partner.

## C-terminal deletions

**[0076]** In an eleventh approach to heterologous expression, the C-terminus of a protein of the invention is mutated. In addition, it is preferred that no fusion partner is used.

[0077] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) the protein's C-terminus region is mutated and, optionally, (b) no fusion partner is used.

**[0078]** The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; manipulating said nucleic acid to mutate nucleotides that encode the protein's C-terminus portion. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector. The first amino acid of the expressed protein will be that of the mature native protein.

[0079] The mutation may be a substitution, insertion or, preferably, a deletion.

**[0080]** This method can increase the levels of expression, particularly for proteins 730, ORF29 and ORF46. For protein 730, a C-terminus region of around 65 to around 214 amino acids may be deleted; for ORF46, the C-terminus region of around 175 amino acids may be deleted; for ORF29, the C-terminus may be deleted to leave around 230-370 N-terminal amino acids.

## Leader peptide mutation

[0081] In a twelfth approach to heterologous expression, the leader peptide of the protein is mutated. This is particularly useful for the expression of protein 919.

[0082] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which the protein's leader peptide is mutated.

**[0083]** The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; and manipulating said nucleic acid to mutate nucleotides within the leader peptide. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector.

## Poly-glycine deletion

**[0084]** In a thirteenth approach to heterologous expression, poly-glycine stretches in wild-type sequences are mutated. This enhances protein expression.

[0085] The poly-glycine stretch has the sequence (Gly)<sub>n</sub>, where n≥4 (e.g. 5, 6, 7, 8, 9 or more). This stretch is mutated to disrupt or remove the (Gly)<sub>n</sub>. This may be by deletion (*e.g.* CGGGGS → CGGGS, CGS or CS), by substitution (*e.g.* CGGGGS → CGXGGS, CGXXGS, CGXGXS *etc.*), and/or by insertion (e.g. CGGGGS → CGGXGGS, CGXGGGS, *etc.*).

**[0086]** This approach is not restricted to Neisserial proteins - it may be used for any protein (particularly bacterial proteins) to enhance heterologous expression. For Neisserial proteins, however, it is particularly suitable for expressing 287, 741, 983 and Tbp2. An alignment of polymorphic forms of 287 is disclosed in WO00/66741.

[0087] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) a poly-glycine stretch within the protein is mutated.

**[0088]** The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; and manipulating said nucleic acid to mutate nucleotides that encode a poly-glycine stretch within the protein sequence. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector.

[0089] Conversely, the opposite approach (i.e. introduction of poly-glycine stretches) can be used to suppress or diminish expression of a given heterologous protein.

### Heterologous host

20

30

35

40

50

55

**[0090]** Whilst expression of the proteins of the invention may take place in the native host (i.e. the organism in which the protein is expressed in nature), the present invention utilises a heterologous host. The heterologous host may be prokaryotic or eukaryotic. It is preferably *E.coli*, but other suitable hosts include *Bacillus subtilis*, *Vibrio cholerae*, *Salmonella typhi*, *Salmonenna typhimurium*, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Neisseria lactamica*, *Neisseria cinerea*, *Mycobateria* (e.g. *M.tuberculosis*), yeast etc.

#### Vectors etc.

[0091] As well as the methods described above, the invention provides (a) nucleic acid and vectors useful in these methods (b) host cells containing said vectors (c) proteins expressed or expressable by the methods (d) compositions comprising these proteins, which may be suitable as vaccines, for instance, or as diagnostic reagents, or as immunogenic compositions (e) these compositions for use as medicaments (e.g. as vaccines) or as diagnostic reagents (f) the use of these compositions in the manufacture of (1) a medicament for treating or preventing infection due to Neisserial bacteria (2) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria, and/or (3) a reagent which can raise antibodies against Neisserial bacteria and (g) a method of treating a patient, comprising administering to the patient a therapeutically effective amount of these compositions.

## Sequences

**[0092]** The invention also provides a protein or a nucleic acid having any of the sequences set out in the following examples. It also provides proteins and nucleic acid having sequence identity to these. As described above, the degree of 'sequence identity' is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more).

[0093] Furthermore, the invention provides nucleic acid which can hybridise to the nucleic acid disclosed in the examples, preferably under "high stringency" conditions (eg. 65°C in a 0.1xSSC, 0.5% SDS solution).

[0094] The invention also provides nucleic acid encoding proteins according to the invention.

**[0095]** It should also be appreciated that the invention provides nucleic acid comprising sequences complementary to those described above (*eq.* for antisense or probing purposes).

**[0096]** Nucleic acid according to the invention can, of course, be prepared in many ways (eg. by chemical synthesis, from genomic or cDNA libraries, from the organism itself etc.) and can take various forms (eg. single stranded, double stranded, vectors, probes etc.).

[0097] In addition, the term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also peptide nucleic acids (PNA) etc.

#### BRIEF DESCRIPTION OF DRAWINGS

[0098]

Figures 1 and 2 show constructs used to express proteins using heterologous leader peptides.

Figure 3 shows expression data for ORF1, and Figure 4 shows similar data for protein 961.

Figure 5 shows domains of protein 287, and Figures 6 & 7 show deletions within domain A.

Figure 8 shows domains of protein 564.

Figure 9 shows the *PhoC* reporter gene driven by the 919 leader peptide, and Figure 10 shows the results obtained using mutants of the leader peptide.

Figure 11 shows insertion mutants of protein 730 (A: 730-C1; B: 730-C2).

Figure 12 shows domains of protein 961.

Figure 13 shows SDS-PAGE of  $\Delta G$  proteins. Dots show the main recombinant product.

Figure 14 shows 26 hybrid proteins according to the invention.

#### MODES FOR CARRYING OUT THE INVENTION

#### Example 1- 919 and its leader peptide

[0099] Protein 919 from *N.meningitidis* (serogroup B, strain 2996) has the following sequence:

```
MKKYLFRAAL YGIAAAILAA CQSKSIQTFP QPDTSVINGP DRPVGIPDPA
51 GTTVGGGGAV YTVVPHLSLP HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV
101 CAQAFQTPVH SFQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDDR
151 RTAQARFPIY GIPDDFISVP LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT
201 HTADLSRFPI TARTTAIKGR FEGSRFLPYH TRNQINGGAL DGKAPILGYA
251 EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301 KLGQTSMQGI KAYMRQNPQR LAEVLGQNPS YIFFRELAGS SNDGPVGALG
351 TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG
401 AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*
```

35

40

55

5

10

15

20

25

30

[0100] The leader peptide is underlined.

[0101] The sequences of 919 from other strains can be found in Figures 7 and 18 of WO00/66741.

[0102] Example 2 of WO99/57280 discloses the expression of protein 919 as a His-fusion in E.coli.

**[0103]** The protein is a good surface-exposed immunogen.

[0104] Three alternative expression strategies were used for 919:

1) 919 without its leader peptide (and without the mature N-terminal cysteine) and without any fusion partner ('919untagged'):

```
45
                        QSKSIQTFP QPDTSVINGP DRPVGIPDPA GTTVGGGGAV YTVVPHLSLP
                    50 HWAAQDFAKS LQSFRLGCAN LKNRQGWQDV CAQAFQTPVH SFQAKQFFER
                   100 YFTPWQVAGN GSLAGTVTGY YEPVLKGDDR RTAQARFPIY GIPDDFISVP
                   150 LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT HTADLSRFPI TARTTAIKGR
                   200
                       FEGSRFLPYH TRNQINGGAL DGKAPILGYA EDPVELFFMH IQGSGRLKTP
50
                       SGKYIRIGYA DKNEHPYVSI GRYMADKGYL KLGQTSMQGI KAYMRQNPQR
                   250
                       LAEVLGQNPS YIFFRELAGS SNDGPVGALG TPLMGEYAGA VDRHYITLGA
                   300
                   350
                       PLFVATAHPV TRKALNRLIM AQDTGSAIKG AVRVDYFWGY GDEAGELAGK
                       QKTTGYVWQL LPNGMKPEYR P*
                   400
```

The leader peptide and cysteine were omitted by designing the 5'-end amplification primer downstream from the predicted leader sequence.

- 2) 919 with its own leader peptide but without any fusion partner ('919L'); and
- 3) 919 with the leader peptide (MKTFFKTLSAAALALILAA) from ORF4 ('919LOrf4').

	1	MKTFFKTLS	AAALALILAA	CQSKSIQTFP	QPDTSVINGP	DRPVGIPDPA	
	50	GTTVGGGGAV	YTVVPHLSLP	HWAAQDFAKS	LQSFRLGCAN	LKNRQGWQDV	
	100	CAQAFQTPVH	SFQAKQFFER	YFTPWQVAGN	GSLAGTVTGY	YEPVLKGDDR	
-	150	RTAQARFPIY	GIPDDFISVP	LPAGLRSGKA	LVRIRQTGKN	SGTIDNTGGT	
5	200	HTADLSRFPI	TARTTAIKGR	FEGSRFLPYH	TRNQINGGAL	DGKAPILGYA	
	250	EDPVELFFMH	IQGSGRLKTP	SGKYIRIGYA	DKNEHPYVSI	GRYMADKGYL	
	300	KLGQTSMQGI	KSYMRQNPQR	LAEVLGQNPS	YIFFRELAGS	SNDGPVGALG	
10							
	350 400				TRKALNRLIM LPNGMKPEYR		

To make this construct, the entire sequence encoding the ORF4 leader peptide was included in the 5'-primer as a tail (primer 919Lorf4 For). A *Nhe*l restriction site was generated by a double nucleotide change in the sequence coding for the ORF4 leader (no amino acid changes), to allow different genes to be fused to the ORF4 leader peptide sequence. A stop codon was included in all the 3'-end primer sequences.

[0105] All three forms of the protein were expressed and could be purified.

**[0106]** The '919L' and '919LOrf4' expression products were both lipidated, as shown by the incorporation of [<sup>3</sup>H]-palmitate label. 919<sup>untagged</sup> did not incorporate the <sup>3</sup>H label and was located intracellularly.

[0107] 919LOrf4 could be purified more easily than 919L. It was purified and used to immunise mice. The resulting sera gave excellent results in FACS and ELISA tests, and also in the bactericidal assay. The lipoprotein was shown to be localised in the outer membrane.

[0108] 919<sup>untagged</sup> gave excellent ELISA titres and high serum bactericidal activity. FACS confirmed its cell surface location.

## Example 2 — 919 and expression temperature

15

30

35

50

**[0109]** Growth of *E.coli* expressing the 919LOrf4 protein at 37°C resulted in lysis of the bacteria. In order to overcome this problem, the recombinant bacteria were grown at 30°C. Lysis was prevented without preventing expression.

### Example 3 - mutation of 907, 919 and 922

[0110] It was hypothesised that proteins 907, 919 and 922 are murein hydrolases, and more particularly lytic transglycosylases. Murein hydrolases are located on the outer membrane and participate in the degradation of peptidoglycan. [0111] The purified proteins 919<sup>untagged</sup>, 919Lorf4, 919-His (*i.e.* with a C-terminus His-tag) and 922-His were thus tested for murein hydrolase activity [Ursinus & Holtje (1994) J.Bact. 176:338-343]. Two different assays were used, one determining the degradation of insoluble murein sacculus into soluble muropeptides and the other measuring breakdown of poly(MurNAc-GlcNAc)<sub>n>30</sub> glycan strands.

[0112] The first assay uses murein sacculi radiolabelled with meso-2,6-diamino-3,4,5-[ $^3$ H]pimelic acid as substrate. Enzyme (3-10  $\mu$ g total) was incubated for 45 minutes at 37°C in a total volume of 100 $\mu$ l comprising 10mM Tris-maleate (pH 5.5), 10mM MgCl<sub>2</sub>, 0.2% v/v Triton X-100 and [ $^3$ H]A<sub>2</sub>pm labelled murein sacculi (about 10000cpm). The assay mixture was placed on ice for 15 minutes with 100  $\mu$ l of 1% w/v N-acetyl-N,N,N-trimethylammonium for 15 minutes and precipitated material pelleted by centrifugation at 10000g for 15 minutes. The radioactivity in the supernatant was measured by liquid scintillation counting. *E.coli* soluble lytic transglycosylase Slt70 was used as a positive control for the assay; the negative control comprised the above assay solution without enzyme.

[0113] All proteins except 919-His gave positive results in the first assay.

[0114] The second assay monitors the hydrolysis of poly(MurNAc-GlcNAc)glycan strands. Purified strands, poly(MurNAc-GlcNAc) $_{n>30}$  labelled with N-acetyl-D-1-[ $^3$ H]glucosamine were incubated with 3 $\mu$ g of 919L in 10 mM Tris-maleate (pH 5.5), 10 mM MgCl $_2$  and 0.2% v/v Triton X-100 for 30 min at 37°C. The reaction was stopped by boiling for 5 minutes and the pH of the sample adjusted to about 3.5 by addition of 10 $\mu$ l of 20% v/v phosphoric acid. Substrate and product were separated by reversed phase HPLC on a Nucleosil 300 C $_{18}$  column as described by Harz et. al. [Anal. Biochem. (1990) 190:120-128]. The *E.coli* lytic transglycosylase Mlt A was used as a positive control in the assay. The negative control was performed in the absence of enzyme.

[0115] By this assay, the ability of 919LOrf4 to hydrolyse isolated glycan strands was demonstrated when anhydrodisaccharide subunits were separated from the oligosaccharide by HPLC.

[0116] Protein 919Lorf4 was chosen for kinetic analyses. The activity of 919Lorf4 was enhanced 3.7-fold by the addition of 0.2% v/v Triton X-100 in the assay buffer. The presence of Triton X-100 had no effect on the activity of 919<sup>untagged</sup>. The effect of pH on enzyme activity was determined in Tris-Maleate buffer over a range of 5.0 to 8.0. The optimal pH for the reaction was determined to be 5.5. Over the temperature range 18°C to 42°C, maximum activity was observed at 37°C. The effect of various ions on murein hydrolase activity was determined by performing the reaction in the presence of a variety of ions at a final concentration of 10mM. Maximum activity was found with Mg²+, which stimulated activity 2.1-fold. Mn²+ and Ca²+ also stimulated enzyme activity to a similar extent while the addition Ni²+ and EDTA had no significant effect. In contrast, both Fe²+and Zn²+ significantly inhibited enzyme activity.

**[0117]** The structures of the reaction products resulting from the digestion of unlabelled *E.coli* murein sacculus were analysed by reversed-phase HPLC as described by Glauner [Anal. Biochem. (1988) 172:451-464]. Murein sacculi digested with the muramidase Cellosyl were used to calibrate and standardise the Hypersil ODS column. The major reaction products were 1,6 anhydrodisaccharide tetra and tri peptides, demonstrating the formation of 1,6 anhydromuraminic acid intramolecular bond.

**[0118]** These results demonstrate experimentally that 919 is a murein hydrolase and in particular a member of the lytic transglycosylase family of enzymes. Furthermore the ability of 922-His to hydrolyse murein sacculi suggests this protein is also a lytic transglycosylase.

[0119] This activity may help to explain the toxic effects of 919 when expressed in E.coli.

**[0120]** In order to eliminate the enzymatic activity, rational mutagenesis was used. 907, 919 and 922 show fairly low homology to three membrane-bound lipidated murein lytic transglycosylases from *E.coli*:

919 (441aa) is 27.3% identical over 440aa overlap to *E.coli* MLTA (P46885);

922 (369aa) is 38.7% identical over 310aa overlap to E.coli MLTB (P41052); and

907-2 (207aa) is 26.8% identical over 149aa overlap to E.coli MLTC (P52066).

907-2 also shares homology with *E.coli* MLTD (P23931) and Slt70 (P03810), a soluble lytic transglycosylase that is located in the periplasmic space. No significant sequence homology can be detected among 919, 922 and 907-2, and the same is true among the corresponding MLTA, MLTB and MLTC proteins.

**[0121]** Crystal structures are available for Slt70 [1QTEA; 1QTEB; Thunnissen et al. (1995) Biochemistry 34: 12729-12737] and for Slt35 [1LTM; 1QUS; 1QUT; van Asselt et al. (1999) Structure Fold Des 7:1167-80] which is a soluble form of the 40kDa MLTB.

[0122] The catalytic residue (a glutamic acid) has been identified for both Slt70 and MLTB.

**[0123]** In the case of Slt70, mutagenesis studies have demonstrated that even a conservative substitution of the catalytic Glu505 with a glutamine (Gln) causes the complete loss of enzymatic activity. Although Slt35 has no obvious sequence similarity to Slt70, their catalytic domains shows a surprising similarity. The corresponding catalytic residue in MLTB is Glu162.

**[0124]** Another residue which is believed to play an important role in the correct folding of the enzymatic cleft is a well-conserved glycine (Gly) downstream of the glutamic acid. Recently, Terrak et al. [Mol.Microbiol. (1999) 34:350-64] have suggested the presence of another important residue which is an aromatic amino acid located around 70-75 residues downstream of the catalytic glutamic acid.

[0125] Sequence alignment of Slt70 with 907-2 and of MLTB with 922 were performed in order to identify the corresponding catalytic residues in the MenB antigens.

[0126] The two alignments in the region of the catalytic domain are reported below:

## 907-2/Slt70:

90 100 110 **▼**120 130 140 907-2.pep ERRRLLVNIQYESSRAG--LDTQIVLGLIEVESAFRQYAISGVGARGLMQVMPFWKNYIG . . :| : :::::: | | | | | : | ||| ||| |||| slty ecoli  $\texttt{ERFPLAYNDLFKRYTSGKEIPQSYAMAIARQ} \textbf{\textit{E}} \texttt{SAWNPKVKSPV} \textbf{\textit{G}} \texttt{ASGLMQIMPGTATHTV}$ ▲ 510 530 480 490 500 520 GLU505

55

50

10

20

30

35

40

## 922/MLTB

		150	160 ▼	170	180	190	200
	922.pep	VAQKYGVPAE	LIVAVIGI <b>E</b> T	ny <b>g</b> knt <b>g</b> sfi	RVADALATLGI	DYPRRAGFF	KELVELLKLA
5		:	:     : :     :	:Ī: Ī: I	:       :	:11111 :1:	: 11 :1 :1
	mltb_ecoli	AWQVYGVPPE	IIVGIIGV <b>E</b> 1	RWGRVMGKTI	RILDALATLSE	NYPRRAEYFS	GELETFLLMA
		150	160 ▲	170	180	190	200
			GI	U162			
		210	220	230	240	250	260
10	922.pep	KEEGGDVFAF	KGSYAGAMGN	ipqfmpss <b>y</b> ri	KWAVDYDGDGH	IRDIWGNVGDV	/AASVANYMKQ
		::   ::	111:11111:	: 1111111::	::   ::	::     :	::     :
	mltb_ecoli	RDEQDDPLNI	KGSFAGAMGY	GQFMPSS <b>Y</b> K(	QYAVDFSGDGH	IINLWDPV-DA	AIGSVANYFKA
		210	220	230	240	250	260

[0127] From these alignments, it results that the corresponding catalytic glutamate in 907-2 is Glu117, whereas in 922 is Glu164. Both antigens also share downstream glycines that could have a structural role in the folding of the enzymatic cleft (in bold), and 922 has a conserved aromatic residue around 70aa downstream (in bold).

**[0128]** In the case of protein 919, no 3D structure is available for its *E.coli* homologue MLTA, and nothing is known about a possible catalytic residue. Nevertheless, three amino acids in 919 are predicted as catalytic residues by alignment with MLTA:

## 919/MLTA

25	919.pep	11: 1 1	1:1::: ::	1:1 : [] []	LKTPSGKYIR: : : ::	: :11 11 1	290 PYVSIGRYMADK           : :   :
	mlta_ecoli.p	ALSDKY-1	LAYSNSLMDN 170	180	190	200	210
30							
35	919.pep		310 SMQGIKSYMR(		LGQNPSYIFF	RELAGSSNDG	
	mlta_ecoli.p						
40	919.pep	360 ▼ EYAGAVDRI	o HYITLGAPLF	380 /ATAHPVTRK	_		00 00410 KGAVRVDYFWGY
	mlta_ecoli.p	: :      RASVASDRS	:  : SIIPPGTTLL	: : : AEVPLLDNNG	:     KFNGQYELRLI	::   : :   MVALDVGGAI	:   :   KGQ-HFDIYQGI
		280	290 20 o	300	310	320	330
45	919.pep	GDEAGELA	GKQKTTGYVW	QLLP			
	mlta_ecoli.p	GPEAGHRAG 340	GWYNHYGRVW 350	/LKT			

- 50 **[0129]** The three possible catalytic residues are shown by the symbol  $\mathbf{\nabla}$ :
  - 1) Glu255 (Asp in MLTA), followed by three conserved glycines (Gly263, Gly265 and Gly272) and three conserved aromatic residues located approximately 75-77 residues downstream. These downstream residues are shown by o.
  - 2) Glu323 (conserved in MLTA), followed by 2 conserved glycines (Gly347 and Gly355) and two conserved aromatic residues located 84-85 residues downstream (Tyr406 or Phe407). These downstream residues are shown by ⋄.
  - 3) Asp362 (instead of the expected Glu), followed by one glycine (Gly 369) and a conserved aromatic residue

(Trp428). These downstream residues are shown by O.

[0130] Alignments of polymorphic forms of 919 are disclosed in WO00/66741.

**[0131]** Based on the prediction of catalytic residues, three mutants of the 919 and one mutant of 907, containing each a single amino acid substitution, have been generated. The glutamic acids in position 255 and 323 and the aspartic acids in position 362 of the 919 protein and the glutamic acid in position 117 of the 907 protein, were replaced with glycine residues using PCR-based SDM. To do this, internal primers containing a codon change from Glu or Asp to Gly were designed:

Primers	Sequences	Codon change				
919-E255 for 919-E255 rev	CGAAGACCCCGTC <u>Ggt</u> CTTTTTTTATG GTGCATAAAAAAAAAGacCGACGGGGTCT	GAA → Ggt				
919-E323 for 919-E323 rev	AACGCCTCGCC <u>Ggt</u> GTTTTGGGTCA TTTGACCCAAAACacCGGCGAGGCG	GAA → Ggt				
919-D362 for 919-D362 rev	TGCCGGCGCAGTC <u>Ggt</u> CGGCACTACA TAATGTAGTGCCGacCGACTGCGCCG	GAC  o Ggt				
907-E117 for 907-E117 rev	TGATTGAGGTG <u>Ggt</u> AGCGCGTTCCG GGCGGAACGCGCTacCCACCTCAAT	GAA  o Ggt				
Underlined nucleotides code for glycine; the mutated nucleotides are in lower case.						

[0132] To generate the 919-E255, 919-E323 and 919-E362 mutants, PCR was performed using 20ng of the pET 919-LOrf4 DNA as template, and the following primer pairs:

1) Orf4L for / 919-E255 rev

10

15

20

25

30

35

40

50

- 2) 919-E255 for / 919L rev
- 3) Orf4L for / 919-E323 rev
- 4) 919-E323 for / 919L rev
- 5) Orf4L for / 919-D362 rev
- 6) 919-D362 for / 919L rev

The second round of PCR was performed using the product of PCR 1-2, 3-4 or 5-6 as template, and as forward and reverse primers the "Orf4L for" and "919L rev" respectively.

For the mutant 907-E117, PCR have been performed using 200ng of chromosomal DNA of the 2996 strain as template and the following primer pairs:

- 7) 907L for / 907-E117 rev
- 8) 907-E 117 for / 907L rev

[0133] The second round of PCR was performed using the products of PCR 7 and 8 as templates and the oligos "907L for" and "907L rev" as primers.

**[0134]** The PCR fragments containing each mutation were processed following the standard procedure, digested with *Ndel* and *Xhol* restriction enzymes and cloned into pET-21b+ vector. The presence of each mutation was confirmed by sequence analysis.

[0135] Mutation of Glu117 to Gly in 907 is carried out similarly, as is mutation of residues Glu164, Ser213 and Asn348 in 922.

**[0136]** The E255G mutant of 919 shows a 50% reduction in activity; the E323G mutant shows a 70% reduction in activity; the E362G mutant shows no reduction in activity.

## Example 4 - multimeric form

[0137] 287-GST, 919<sup>untagged</sup> and 953-His were subjected to gel filtration for analysis of quaternary structure or preparative purposes. The molecular weight of the native proteins was estimated using either FPLC Superose 12 (H/R 10/30) or Superdex 75 gel filtration columns (Pharmacia). The buffers used for chromatography for 287, 919 and 953 were 50 mM Tris-HCl (pH 8.0), 20 mM Bicine (pH 8.5) and 50 mM Bicine (pH 8.0), respectively.

[0138] Additionally each buffer contained 150-200 mM NaCl and 10% v/v glycerol. Proteins were dialysed against the appropriate buffer and applied in a volume of 200 µl. Gel filtration was performed with a flow rate of 0.5 - 2.0 ml/min and

the eluate monitored at 280nm. Fractions were collected and analysed by SDS-PAGE. Blue dextran 2000 and the molecular weight standards ribonuclease A, chymotrypsin A ovalbumin, albumin (Pharmacia) were used to calibrate the column. The molecular weight of the sample was estimated from a calibration curve of K<sub>av</sub> vs. log M<sub>r</sub> of the standards. Before gel filtration, 287-GST was digested with thrombin to cleave the GST moiety.

**[0139]** The estimated molecular weights for 287, 919 and 953-His were 73 kDa, 47 kDa and 43 kDa respectively. These results suggest 919 is monomeric while both 287 and 953 are principally dimeric in their nature. In the case of 953-His, two peaks were observed during gel filtration. The major peak (80%) represented a dimeric conformation of 953 while the minor peak (20%) had the expected size of a monomer. The monomeric form of 953 was found to have greater bactericidal activity than the dimer.

## Example 5 - pSM214 and pET-24b vectors

10

30

35

40

45

50

55

[0140] 953 protein with its native leader peptide and no fusion partners was expressed from the pET vector and also from pSM214 [Velati Bellini et al. (1991) J. Biotechnol. 18, 177-192].

[0141] The 953 sequence was cloned as a full-length gene into pSM214 using the *E. coli* MM294-1 strain as a host. To do this, the entire DNA sequence of the 953 gene (from ATG to the STOP codon) was amplified by PCR using the following primers:

953L for/2 CCGGAATTCTTATGAAAAAAATCATCTTCGCCGC Eco RI 953L rev/2 GCCAAGCTTTTATTGTTTGGCTGCCTCGATT Hind III

which contain *Eco*RI and *Hind*III restriction sites, respectively. The amplified fragment was digested with *Eco*RI and *Hind*III and ligated with the pSM214 vector digested with the same two enzymes. The ligated plasmid was transformed into *E.coli* MM294-1 cells (by incubation in ice for 65 minutes at 37° C) and bacterial cells plated on LB agar containing 20μg/ml of chloramphenicol.

[0142] Recombinant colonies were grown over-night at  $37^{\circ}$ C in 4 ml of LB broth containing 20  $\mu$ g/ml of chloramphenicol; bacterial cells were centrifuged and plasmid DNA extracted as and analysed by restriction with *Eco*Rl and *Hind*III. To analyse the ability of the recombinant colonies to express the protein, they were inoculated in LB broth containing  $20\mu$ g/ml of chloramphenicol and let to grown for 16 hours at  $37^{\circ}$ C. Bacterial cells were centrifuged and resuspended in PBS. Expression of the protein was analysed by SDS-PAGE and Coomassie Blue staining.

[0143] Expression levels were unexpectedly high from the pSM214 plasmid.

[0144] Oligos used to clone sequences into pSM-214 vectors were as follows:

∆ <b>G287</b>	Fwd	CCG <u>GAATTC</u> TTATG-TCGCCCGATGTTAAATCGGCGGA	EcoRI
(pSM-214)	Rev	GCCCAAGCTT-TCAATCCTGCTCTTTTTTGCCG	HindIII
∆2 287	Fwd	CCGGAATTCTTATG-AGCCAAGATATGGCGGCAGT	EcoRI
(pSM-214)	Rev	GCCCAAGCTT-TCAATCCTGCTCTTTTTTGCCG	HindIII
∆3 287	Fwd	CCG <u>GAATTC</u> TTATG-TCCGCCGAATCCGCAAATCA	EcoRI
(pSM-214)	Rev	GCCCAAGCTT-TCAATCCTGCTCTTTTTTGCCG	HindIII
∆4 287	Fwd	CCG <u>GAATTC</u> TTATG-GGAAGGGTTGATTTGGCTAATG	EcoRI
(pSM-214)	Rev	GCCCAAGCTT-TCAATCCTGCTCTTTTTTGCCG	HindIII
Orf46.1	Fwd	CCG <u>GAATTC</u> TTATG-TCAGATTTGGCAAACGATTCTT	EcoRI
(pSM-214)	Rev	GCCC <u>AAGCTT</u> - <b>TTA</b> CGTATCATATTTCACGTGCTTC	HindIII
∆G287-Orf46.1	Fwd	CCG <u>GAATTC</u> TTATG-TCGCCCGATGTTAAATCGGCGGA	EcoRI
(pSM-214)	Rev	GCCCAAGCTT-TTACGTATCATATTTCACGTGCTTC	HindIII
919	Fwd	CCG <u>GAATTC</u> TTATG-CAAAGCAAGAGCATCCAAACCT	EcoRI
(pSM-214)	Rev	GCCCAAGCTT-TTACGGGCGGTATTCGGGCT	HindIII
961L	Fwd	CCG <u>GAATTC</u> ATATG-AAACACTTTCCATCC	EcoRI
(pSM-214)	Rev	GCCC <u>AAGCTT</u> - <b>TTA</b> CCACTCGTAATTGAC	HindIII
961	Fwd	CCG <u>GAATTC</u> ATATG-GCCACAAGCGACGAC	EcoRI
(pSM-214)	Rev	GCCC <u>AAGCTT</u> - <b>TTA</b> CCACTCGTAATTGAC	HindIII

(continued)

961c L	Fwd	CCG <u>GAATTC</u> TTATG-AAACACTTTCCATCC	EcoRI
pSM-214	Rev	GCCCAAGCTT-TCAACCCACGTTGTAAGGTTG	HindIII
961c	Fwd	CCG <u>GAATTC</u> TTATG-GCCACAAACGACGACG	EcoRI
pSM-214	Rev	GCCCAAGCTT-TCAACCCACGTTGTAAGGTTG	HindIII
953	Fwd	CCG <u>GAATTC</u> TTATG-GCCACCTACAAAGTGGACGA	EcoRI
(pSM-214)	Rev	GCCC <u>AAGCTT</u> -TTATTGTTTGGCTGCCTCGATT	HindIII

These sequences were manipulated, cloned and expressed as described for 953L.

**[0145]** For the pET-24 vector, sequences were cloned and the proteins expressed in pET-24 as described below for pET21. pET2 has the same sequence as pET-21, but with the kanamycin resistance cassette instead of ampicillin cassette.

[0146] Oligonucleotides used to clone sequences into pET-24b vector were:

∆G 287 K	Fwd	CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC §	Nhel
	Rev	CCCG <u>CTCGAG</u> - <b>TCA</b> ATCCTGCTCTTTTTTGCC *	Xhol
∆2 287 K	Fwd	CGCGGATCCGCTAGC-CAAGATATGGCGGCAGT §	
∆3 287 K	Fwd	CGCGGATCC <u>GCTAGC</u> -GCCGAATCCGCAAATCA §	Nhel
∆4 287 K	4 287 K Fwd CGCGCTAGC-GGAAGGGTTGATTTGGCTAATGG §		Nhel
Orf46.1 K	Fwd	GGGAATTC <u>CATATG</u> -GGCATTTCCCGCAAAATATC	Ndel
	Rev	CCCG <u>CTCGAG</u> - <b>TTA</b> CGTATCATATTTCACGTGC	Xhol
Orf46A K	Fwd	GGGAATTCCATATG-GGCATTTCCCGCAAAATATC	Ndel
	Rev	CCCG <u>CTCGAG</u> - <b>TTA</b> TTCTATGCCTTGTGCGGCAT	Xhol
961 K	Fwd	CGCGGATCC <u>CATATG</u> -GCCACAAGCGACGACGA	Ndel
(MC58)	Rev	CCCG <u>CTCGAG</u> - <b>TTA</b> CCACTCGTAATTGAC	Xhol
961a K	Fwd	CGCGGATCC <u>CATATG</u> -GCCACAAACGACG	Ndel
	Rev	CCCG <u>CTCGAG</u> - <b>TCA</b> TTTAGCAATATTATCTTTGTTC	Xhol
961b K	Fwd	CGCGGATCC <u>CATATG</u> -AAAGCAAACAGTGCCGAC	Ndel
	Rev	CCCG <u>CTCGAG</u> - <b>TTA</b> CCACTCGTAATTGAC	Xhol
961c K	Fwd	CGCGGATCC <u>CATATG</u> -GCCACAAACGACG	Ndel
	Rev	CCCG <u>CTCGAG</u> - <b>TTA</b> ACCCACGTTGTAAGGT	Xhol
961cL K	Fwd	CGCGGATCC <u>CATATG</u> -ATGAAACACTTTCCATCC	Ndel
	Rev	CCCG <u>CTCGAG</u> - <b>TTA</b> ACCCACGTTGTAAGGT	Xhol
961d K	Fwd	CGCGGATCC <u>CATATG</u> -GCCACAAACGACG	Ndel
	Rev	CCCGCTCGAG- <b>TCA</b> GTCTGACACTGTTTTATCC	Xhol
∆G 287-	Fwd	CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC	Nhel
919 K	Rev	CCCG <u>CTCGAG</u> -TTACGGGCGGTATTCGG	Xhol
∆G 287-	Fwd	CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC	Nhel
Orf46.1 K	Rev	CCCG <u>CTCGAG</u> - <b>TTA</b> CGTATCATATTTCACGTGC	Xhol
∆G 287-	Fwd	CGCGGATCC <u>GCTAGC</u> -CCCGATGTTAAATCGGC	Nhel

(continued)

961 K	Rev	CCCG <u>CTCGAG</u> - <b>TTA</b> CCACTCGTAATTGAC	Xhol
,		sed as a Reverse primer for all the 287 forms. sed in combination with the ΔG278 K reverse primer.	

# Example 6 - ORF1 and its leader peptide

5

40

45

50

55

[0147] ORF1 from *N.meningitidis* (serogroup B, strain MC58) is predicted to be an outer membrane or secreted protein. It has the following sequence:

	1	MKTTDKRTTE	THRKAPKTGR	IRFSPAYLAI	CLSFGILPQA	WAGHTYFGIN
	51	YQYYRDFAEN	KGKFAVGAKD	IEVYNKKGEL	VGKSMTKAPM	IDFSVVSRNG
15	101	VAALVGDQYI	VSVAHNGGYN	NVDFGAEGRN	PDQHRFTYKI	VKRNNYKAGT
	151	KGHPYGGDYH	MPRLHKFVTD	AEPVEMTSYM	DGRKYIDQNN	YPDRVRIGAG
	201	RQYWRSDEDE	PNNRESSYHI	ASAYSWLVGG	NTFAQNGSGG	GTVNLGSEKI
	251	KHSPYGFLPT	GGSFGDSGSP	MFIYDAQKQK	WLINGVLQTG	NPYIGKSNGF
	301	QLVRKDWFYD	EIFAGDTHSV	FYEPRQNGKY	SFNDDNNGTG	KINAKHEHNS
	351	LPNRLKTRTV	QLFNVSLSET	AREPVYHAAG	GVNSYRPRLN	NGENISFIDE
20	401	GKGELILTSN	INQGAGGLYF	QGDFTVSPEN	NETWQGAGVH	ISEDSTVTWK
	451	VNGVANDRLS	KIGKGTLHVQ	AKGENQGSIS	VGDGTVILDQ	QADDKGKKQA
	501	FSEIGLVSGR	GTVQLNADNQ	FNPDKLYFGF	RGGRLDLNGH	SLSFHRIQNT
	551	DEGAMIVNHN	QDKESTVTIT	GNKDIATTGN	NNSLDSKKEI	AYNGWFGEKD
	601	TTKTNGRLNL	VYQPAAEDRT	LLLSGGTNLN	GNITQTNGKL	FFSGRPTPHA
25	651	YNHLNDHWSQ	KEGIPRGEIV	WDNDWINRTF	KAENFQIKGG	QAVVSRNVAK
25	701	VKGDWHLSNH	AQAVFGVAPH	QSHTICTRSD	WTGLTNCVEK	TITDDKVIAS
	751	LTKTDISGNV	DLADHAHLNL	TGLATLNGNL	SANGDTRYTV	SHNATQNGNL
	801	SLVGNAQATF	NQATLNGNTS	ASGNASFNLS	DHAVQNGSLT	LSGNAKANVS
	851	HSALNGNVSL	ADKAVFHFES	SRFTGQISGG	KDTALHLKDS	EWTLPSGTEL
	901	GNLNLDNATI	TLNSAYRHDA	AGAQTGSATD	APRRRSRRSR	RSLLSVTPPT
30	951	SVESRFNTLT	VNGKLNGQGT	FRFMSELFGY	RSDKLKLAES	SEGTYTLAVN
	1001	NTGNEPASLE	QLTVVEGKDN	KPLSENLNFT	LQNEHVDAGA	WRYQLIRKDG
	1051	EFRLHNPVKE	QELSDKLGKA	EAKKQAEKDN	AQSLDALIAA	GRDAVEKTES
	1101	VAEPARQAGG	ENVGIMQAEE	EKKRVQADKD	TALAKQREAE	TRPATTAFPR
	1151	ARRARRDLPQ	LQPQPQPQPQ	RDLISRYANS	GLSEFSATLN	SVFAVQDELD
	1201	RVFAEDRRNA	VWTSGIRDTK	HYRSQDFRAY	RQQTDLRQIG	MQKNLGSGRV
35	1251	GILFSHNRTE	NTFDDGIGNS	ARLAHGAVFG	QYGIDRFYIG	ISAGAGFSSG
	1301	SLSDGIGGKI	RRRVLHYGIQ	ARYRAGFGGF	GIEPHIGATR	YFVQKADYRY
	1351	ENVNIATPGL	AFNRYRAGIK	ADYSFKPAQH	ISITPYLSLS	YTDAASGKVR
	1401	TRVNTAVLAQ	DFGKTRSAEW	GVNAEIKGFT	LSLHAAAAKG	PQLEAQHSAG
	1451	IKLGYRW*				

The leader peptide is underlined.

[0148] A polymorphic form of ORF1 is disclosed in WO99/55873.

[0149] Three expression strategies have been used for ORF1:

- 1) ORF1 using a His tag, following WO99/24578 (ORF1-His);
- 2) ORF1 with its own leader peptide but without any fusion partner ('ORF1L'); and
- 3) ORF1 with the leader peptide (MKKTAIAIAVALAGFATVAQA) from *E.coli* OmpA ('Orf1LOmpA'):

MKKTAIAIAVALAGFATVAQAASAGHTYFGINYQYYRDFAENKGKFAVGAKDIEVYNKKGELVGKSMTKAPMIDFSV VSRNGVAALVGDQYIVSVAHNGGYNNVDFGAEGRNPDQHRFTYKIVKRNNYKAGTKGHPYGGDYHMPRLHKFVTDAE PVEMTSYMDGRKYIDQNNYPDRVRIGAGRQYWRSDEDEPNNRESSYHIASAYSWLVGGNTFAQNGSGGGTVNLGSEK IKHSPYGFLPTGGSFGDSGSPMFIYDAQKQKWLINGVLQTGNPYIGKSNGFQLVRKDWFYDEIFAGDTHSVFYEPRQ NGKYSFNDDNNGTGKINAKHEHNSLPNRLKTRTVQLFNVSLSETAREPVYHAAGGVNSYRPRLNNGENISFIDEGKG  $\verb|ELILTSNINQGAGGLYFQGDFTVSPENNETWQGAGVHISEDSTVTWKVNGVANDRLSKIGKGTLHVQAKGENQGSIS||$ VGDGTVILDQQADDKGKKQAFSEIGLVSGRGTVQLNADNQFNPDKLYFGFRGGRLDLNGHSLSFHRIQNTDEGAMIV NHNQDKESTVTITGNKDIATTGNNNSLDSKKEIAYNGWFGEKDTTKTNGRLNLVYQPAAEDRTLLLSGGTNLNGNIT QTNGKLFFSGRPTPHAYNHLNDHWSQKEGIPRGEIVWDNDWINRTFKAENFQIKGGQAVVSRNVAKVKGDWHLSNHA QAVFGVAPHQSHTICTRSDWTGLTNCVEKTITDDKVIASLTKTDISGNVDLADHAHLNLTGLATLNGNLSANGDTRY TVSHNATQNGNLSLVGNAQATFNQATLNGNTSASGNASFNLSDHAVQNGSLTLSGNAKANVSHSALNGNVSLADKAV FHFESSRFTGQISGGKDTALHLKDSEWTLPSGTELGNLNLDNATITLNSAYRHDAAGAQTGSATDAPRRRSRRSRRS LLSVTPPTSVESRFNTLTVNGKLNGQGTFRFMSELFGYRSDKLKLAESSEGTYTLAVNNTGNEPASLEQLTVVEGKD NKPLSENLNFTLQNEHVDAGAWRYQLIRKDGEFRLHNPVKEQELSDKLGKAEAKKQAEKDNAQSLDALIAAGRDAVE KTESVAEPARQAGGENVGIMQAEEEKKRVQADKDTALAKQREAETRPATTAFPRARRARRDLPQLQPQPQPQPQPQRDL ISRYANSGLSEFSATLNSVFAVODELDRVFAEDRRNAVWTSGIRDTKHYRSQDFRAYRQQTDLRQIGMQKNLGSGRV GILFSHNRTENTFDDGIGNSARLAHGAVFGOYGIDRFYIGISAGAGFSSGSLSDGIGGKIRRRVLHYGIQARYRAGF GGFGIEPHIGATRYFVQKADYRYENVNIATPGLAFNRYRAGIKADYSFKPAQHISITPYLSLSYTDAASGKVRTRVN TAVLAQDFGKTRSAEWGVNAEIKGFTLSLHAAAAKGPQLEAQHSAGIKLGYRW\*

20

25

40

5

10

15

To make this construct, the clone pET911 LOmpA (see below) was digested with the *Nhel* and *Xhol* restriction enzymes and the fragment corresponding to the vector carrying the OmpA leader sequence was purified (pETLOmpA). The ORF1 gene coding for the mature protein was amplified using the oligonucleotides ORF1-For and ORF1-Rev (including the *Nhel* and *Xhol* restriction sites, respectively), digested with *Nhel* and *Xhol* and ligated to the purified pETOmpA fragment (see Figure 1). An additional AS dipeptide was introduced by the *Nhel* site.

**[0150]** All three forms of the protein were expressed. The His-tagged protein could be purified and was confirmed as surface exposed, and possibly secreted (see Figure 3). The protein was used to immunise mice, and the resulting sera gave excellent results in the bactericidal assay.

[0151] ORF1LOmpA was purified as total membranes, and was localised in both the inner and outer membranes. Unexpectedly, sera raised against ORF1LOmpA show even better ELISA and anti-bactericidal properties than those raised against the His-tagged protein.

[0152] ORF1L was purified as outer membranes, where it is localised.

## 35 Example 7 - protein 911 and its leader peptide

[0153] Protein 911 from *N.meningitidis* (serogroup B, strain MC58) has the following sequence:

- 1 MKKNILEFWV GLFVLIGAAA VAFLAFRVAG GAAFGGSDKT YAVYADFGDI
- 51 GGLKVNAPVK SAGVLVGRVG AIGLDPKSYQ ARVRLDLDGK YQFSSDVSAQ
- 101 ILTSGLLGEQ YIGLQQGGDT ENLAAGDTIS VTSSAMVLEN LIGKFMTSFA
- 151 EKNADGGNAE KAAE\*
- The leader peptide is underlined.

[0154] Three expression strategies have been used for 911:

- 1) 911 with its own leader peptide but without any fusion partner ('911L');
- 2) 911 with the leader peptide from *E.coli* OmpA ('911LOmpA'). To make this construct, the entire sequence encoding the OmpA leader peptide was included in the 5'- primer as a tail (primer 911LOmpA Forward). A *Nhel* restriction site was inserted between the sequence coding for the OmpA leader peptide and the 911 gene encoding the predicted mature protein (insertion of one amino acid, a serine), to allow the use of this construct to clone different genes downstream the OmpA leader peptide sequence.
- 3) 911 with the leader peptide (MKYLLPTAAAGLLLAAQPAMA) from Erwinia carotovora PelB ('911LpelB').

55

50

**[0155]** To make this construct, the 5'-end PCR primer was designed downstream from the leader sequence and included the *Ncol* restriction site in order to have the 911 fused directly to the PelB leader sequence; the 3'- end primer included the STOP codon. The expression vector used was pET22b+ (Novagen), which carries the coding sequence

for the PelB leader peptide. The Ncol site introduces an additional methionine after the PelB sequence.

[0156] All three forms of the protein were expressed. ELISA titres were highest using 911 L, with 919LOmpA also giving good results.

## Example 8 - ORF46

[0157] The complete ORF46 protein from *N.meningitidis* (serogroup B, strain 2996) has the following sequence:

10	1	LGISRKISLI	LSILAVCLPM	<u>HAHA</u> SDLAND	SFIRQVLDRQ	HFEPDGKYHL
	51	FGSRGELAER	SGHIGLGKIQ	SHQLGNLMIQ	QAAIKGNIGY	IVRFSDHGHE
	101	VHSPFDNHAS	HSDSDEAGSP	VDGFSLYRIH	WDGYEHHPAD	GYDGPQGGGY
	151	PAPKGARDIY	SYDIKGVAQN	IRLNLTDNRS	TGQRLADRFH	NAGSMLTQGV
	201	GDGFKRATRY	SPELDRSGNA	AEAFNGTADI	VKNIIGAAGE	IVGAGDAVQG
	251	ISEGSNIAVM	HGLGLLSTEN	KMARINDLAD	MAQLKDYAAA	AIRDWAVQNP
15	301	NAAQGIEAVS	NIFMAAIPIK	GIGAVRGKYG	LGGITAHPIK	RSQMGAIALP
	351	KGKSAVSDNF	ADAAYAKYPS	PYHSRNIRSN	LEQRYGKENI	TSSTVPPSNG
	401	KNVKLADQRH	PKTGVPFDGK	GFPNFEKHVK	YDTKLDIQEL	SGGGIPKAKP
	451	VSDAKPRWEV	DRKLNKLTTR	EQVEKNVQEI	RNGNKNSNFS	QHAQLEREIN
	501	KLKSADEINF	ADGMGKFTDS	MNDKAFSRLV	KSVKENGFTN	PVVEYVEING
	551	KAYIVRGNNR	VFAAEYLGRI	HELKFKKVDF	PVPNTSWKNP	TDVLNESGNV
20	601	KRPRYRSK*				

The leader peptide is underlined.

[0158] The sequences of ORF46 from other strains can be found in WO00/66741.

[0159] Three expression strategies have been used for ORF46:

- 1) ORF46 with its own leader peptide but without any fusion partner ('ORF46-2L');
- 2) ORF46 without its leader peptide and without any fusion partner ('ORF46-2'), with the leader peptide omitted by designing the 5'-end amplification primer downstream from the predicted leader sequence:

	1 SE	LANDSFIR	QVLDRQHFEP	DGKYHLFGSR	GELAERSGHI	GLGKIQSHQL
	51 GN	ILMIQQAAI	KGNIGYIVRF	SDHGHEVHSP	FDNHASHSDS	DEAGSPVDGF
1	01 SI	YRIHWDGY	EHHPADGYDG	PQGGGYPAPK	GARDIYSYDI	KGVAQNIRLN
35	51 LT	DNRSTGQR	LADRFHNAGS	MLTQGVGDGF	KRATRYSPEL	DRSGNAAEAF
2	01 NG	TADIVKNI	IGAAGEIVGA	GDAVQGISEG	SNIAVMHGLG	LLSTENKMAR
2	51 IN	IDLADMAQL	KDYAAAAIRD	WAVQNPNAAQ	GIEAVSNIFM	AAIPIKGIGA
3	01 VR	RGKYGLGGI	TAHPIKRSQM	GAIALPKGKS	AVSDNFADAA	YAKYPSPYHS
3.	51 RN	IIRSNLEQR	YGKENITSST	VPPSNGKNVK	LADQRHPKTG	VPFDGKGFPN
4	01 FE	KHVKYDTK	LDIQELSGGG	IPKAKPVSDA	KPRWEVDRKL	NKLTTREQVE
40 4.	51 KN	IVQEIRNGN	KNSNFSQHAQ	LEREINKLKS	ADEINFADGM	GKFTDSMNDK
5	01 AF	SRLVKSVK	ENGFTNPVVE	YVEINGKAYI	VRGNNRVFAA	EYLGRIHELK
5	51 FK	KVDFPVPN	${\tt TSWKNPTDVL}$	NESGNVKRPR	YRSK*	

3) ORF46 as a truncated protein, consisting of the first 433 amino acids ('ORF46.1L'), constructed by designing PCR primers to amplify a partial sequence corresponding to aa 1-433.

A STOP codon was included in the 3'-end primer sequences.

**[0160]** ORF46-2L is expressed at a very low level to *E. coli*. Removal of its leader peptide (ORF46-2) does not solve this problem. The truncated ORF46.1L form (first 433 amino acids, which are well conserved between serogroups and species), however, is well-expressed and gives excellent results in ELISA test and in the bactericidal assay.

**[0161]** ORF46.1 has also been used as the basis of hybrid proteins. It has been fused with 287, 919, and ORF1. The hybrid proteins were generally insoluble, but gave some good ELISA and bactericidal results (against the homologous 2996 strain):

Protein	ELISA	Bactericidal Ab
Orf1-Orf46.1-His	850	256

55

45

50

(continued)

Protein	ELISA	Bactericidal Ab
919-Orf46.1-His	12900	512
919-287-Orf46-His	n.d.	n.d.
Orf46.1-287His	150	8192
Orf46.1-919His	2800	2048
Orf46.1-287-919His	3200	16384

**[0162]** For comparison, 'triple' hybrids of ORF46.1, 287 (either as a GST fusion, or in  $\Delta$ G287 form) and 919 were constructed and tested against various strains (including the homologous 2996 strain) *versus* a simple mixture of the three antigens. FCA was used as adjuvant:

	2996	BZ232	MC58	NGH38	F6124	BZ133
Mixture	8192	256	512	1024	>2048	>2048
ORF46.1-287-919his	16384	256	4096	8192	8192	8192
∆G287-919-ORF46.1his	8192	64	4096	8192	8192	16384
∆G287-ORF46.1-919his	4096	128	256	8192	512	1024

Again, the hybrids show equivalent or superior immunological activity.

[0163] Hybrids of two proteins (strain 2996) were compared to the individual proteins against various heterologous strains:

	1000	MC58	F6124 (MenA)
ORF46.1-His	<4	4096	<4
ORF1-His	8	256	128
ORF1—ORF46.1-His	1024	512	1024

[0164] Again, the hybrid shows equivalent or superior immunological activity.

## Example 9 - protein 961

5

10

15

20

25

30

35

40

45

50

55

[0165] The complete 961 protein from *N.meningitidis* (serogroup B, strain MC58) has the following sequence:

1	MSMKHFPAKV	LTTAILATFC	SGALAATSDD	DVKKAATVAI	VAAYNNGQEI
51	NGFKAGETIY	DIGEDGTITQ	KDATAADVEA	DDFKGLGLKK	VVTNLTKTVN
101	ENKQNVDAKV	KAAESEIEKL	TTKLADTDAA	LADTDAALDE	TTNALNKLGE
151	NITTFAEETK	TNIVKIDEKL	EAVADTVDKH	AEAFNDIADS	LDETNTKADE
201	AVKTANEAKQ	TAEETKQNVD	AKVKAAETAA	GKAEAAAGTA	NTAADKAEAV
251	AAKVTDIKAD	IATNKADIAK	NSARIDSLDK	NVANLRKETR	QGLAEQAALS
301	GLFQPYNVGR	FNVTAAVGGY	KSESAVAIGT	GFRFTENFAA	KAGVAVGTSS
351	GSSAAYHVGV	NYEW*			

[0166] The leader peptide is underlined.

[0167] Three approaches to 961 expression were used:

- 1) 961 using a GST fusion, following WO99/57280 ('GST961');
- 2) 961 with its own leader peptide but without any fusion partner ('961L'); and
- 3) 961 without its leader peptide and without any fusion partner ('961untagged,), with the leader peptide omitted by

designing the 5'-end PCR primer downstream from the predicted leader sequence.

[0168] All three forms of the protein were expressed. The GST-fusion protein could be purified and antibodies against it confirmed that 961 is surface exposed (Figure 4). The protein was used to immunise mice, and the resulting sera gave excellent results in the bactericidal assay. 961 L could also be purified and gave very high ELISA titres.

[0169] Protein 961 appears to be phase variable. Furthermore, it is not found in all strains of N.meningitidis.

## Example 10 - protein 287

10

30

35

40

45

50

55

[0170] Protein 287 from *N.meningitidis* (serogroup B, strain 2996) has the following sequence:

```
1 MFERSVIAMA CIFALSACGG GGGGSPDVKS ADTLSKPAAP VVAEKETEVK
51 EDAPQAGSQG QGAPSTQGSQ DMAAVSAENT GNGGAATTDK PKNEDEGPQN
101 DMPQNSAESA NQTGNNQPAD SSDSAPASNP APANGGSNFG RVDLANGVLI
151 DGPSQNITLT HCKGDSCNGD NLLDEEAPSK SEFENLNESE RIEKYKKDGK

201 SDKFTNLVAT AVQANGTNKY VIIYKDKSAS SSSARFRRSA RSRRSLPAEM
251 PLIPVNQADT LIVDGEAVSL TGHSGNIFAP EGNYRYLTYG AEKLPGGSYA
301 LRVQGEPAKG EMLAGTAVYN GEVLHFHTEN GRPYPTRGRF AAKVDFGSKS
351 VDGIIDSGDD LHMGTQKFKA AIDGNGFKGT WTENGGGDVS GRFYGPAGEE
401 VAGKYSYRPT DAEKGGFGVF AGKKEQD*
```

[0171] The leader peptide is shown underlined.

[0172] The sequences of 287 from other strains can be found in Figures 5 and 15 of WO00/66741.

[0173] Example 9 of WO99/57280 discloses the expression of 287 as a GST-fusion in E.coli.

[0174] A number of further approaches to expressing 287 in E.coli have been used, including:

- 1) 287 as a His-tagged fusion ('287-His');
- 2) 287 with its own leader peptide but without any fusion partner ('287L');
- 3) 287 with the ORF4 leader peptide and without any fusion partner ('287LOrf4'); and
- 4) 287 without its leader peptide and without any fusion partner ('287untagged'):

```
CGGGGGGSPD VKSADTLSKP AAPVVAEKET EVKEDAPQAG SQGQAPSTQ 51 GSQDMAAVSA ENTGNGGAAT TDKPKNEDEG PQNDMPQNSA ESANQTGNNQ 101 PADSSDSAPA SNPAPANGGS NFGRVDLANG VLIDGPSQNI TLTHCKGDSC 151 NGDNLLDEEA PSKSEFENLN ESERIEKYKK DGKSDKFTNL VATAVQANGT 201 NKYVIIYKDK SASSSSAFFR RSARSRRSLP AEMPLIPVNQ ADTLIVDGEA 251 VSLTGHSGNI FAPEGNYRYL TYGAEKLPGG SYALRVQGEP AKGEMLAGTA 301 VYNGEVLHFH TENGRPYPTR GRFAAKVDFG SKSVDGIIDS GDDLHMGTQK 351 FKAAIDGNGF KGTWTENGGG DVSGRFYGPA GEEVAGKYSY RPTDAEKGGF 401 GVFAGKKEQD *
```

[0175] All these proteins could be expressed and purified. [0176] '287L' and '287LOrf4' were confirmed as lipoproteins.

[0177] As shown in Figure 2, '287LOrf4' was constructed by digesting 919LOrf4 with *Nhel* and *Xhol*. The entire ORF4 leader peptide was restored by the addition of a DNA sequence coding for the missing amino acids, as a tail, in the 5'-end primer (287LOrf4 for), fused to 287 coding sequence. The 287 gene coding for the mature protein was amplified using the oligonucleotides 287LOrf4 For and Rev (including the *Nhel* and *Xhol* sites, respectively), digested with *Nhel* and *Xhol* and ligated to the purified pETOrf4 fragment.

## Example 11 - further non-fusion proteins with/without native leader peptides

[0178] A similar approach was adopted for *E.coli* expression of further proteins from WO99/24578, WO99/36544 and WO99/57280.

**[0179]** The following were expressed without a fusion partner: 008, 105, 117-1, 121-1, 122-1, 128-1, 148, 216, 243, 308, 593, 652, 726, 982, and Orf143-1. Protein 117-1 was confirmed as surface-exposed by FACS and gave high ELISA titres.

[0180] The following were expressed with the native leader peptide but without a fusion partner: 111, 149, 206, 225-1, 235, 247-1, 274, 283, 286, 292, 401, 406, 502-1, 503, 519-1, 525-1, 552, 556, 557, 570, 576-1, 580, 583, 664, 759, 907, 913, 920-1, 926, 936-1, 953, 961, 983, 989, Orf4, Orf7-1, Orf9-1, Orf23, Orf25, Orf37, Orf38, Orf40, Orf40.1, Orf40.2, Orf72-1, Orf76-1, Orf85-2, Orf91, Orf97-1, Orf119, Orf143.1. These proteins are given the suffix 'L'.

[0181] His-tagged protein 760 was expressed with and without its leader peptide. The deletion of the signal peptide greatly increased expression levels. The protein could be purified most easily using 2M urea for solubilisation.

[0182] His-tagged protein 264 was well-expressed using its own signal peptide, and the 30kDa protein gave positive Western blot results.

[0183] All proteins were successfully expressed.

[0184] The localisation of 593, 121-1, 128-1, 593, 726, and 982 in the cytoplasm was confirmed.

[0185] The localisation of 920-1L, 953L, ORF9-1L, ORF85-2L, ORF97-1L, 570L, 580L and 664L in the periplasm was confirmed.

[0186] The localisation of ORF40L in the outer membrane, and 008 and 519-1L in the inner membrane was confirmed. ORF25L, ORF4L, 406L, 576-1L were all confirmed as being localised in the membrane.

[0187] Protein 206 was found not to be a lipoprotein.

**[0188]** ORF25 and ORF40 expressed with their native leader peptides but without fusion partners, and protein 593 expressed without its native leader peptide and without a fusion partner, raised good anti-bactericidal sera. Surprisingly, the forms of ORF25 and ORF40 expressed without fusion partners and using their own leader peptides (i.e. 'ORF25L' and 'ORF40L') give better results in the bactericidal assay than the fusion proteins.

**[0189]** Proteins 920L and 953L were subjected to N-terminal sequencing, giving HRVWVETAH and ATYKVDEY-HANARFAF, respectively. This sequencing confirms that the predicted leader peptides were cleaved and, when combined with the periplasmic location, confirms that the proteins are correctly processed and localised by *E.coli* when expressed from their native leader peptides.

**[0190]** The N-terminal sequence of protein 519.1L localised in the inner membrane was MEFFIILLA, indicating that the leader sequence is not cleaved. It may therefore function as both an uncleaved leader sequence and a transmembrane anchor in a manner similar to the leader peptide of PBP1 from *N.gonorrhoeae* [Ropp & Nicholas (1997) J. Bact. 179: 2783-2787.]. Indeed the N-terminal region exhibits strong hydrophobic character and is predicted by the Tmpred. program to be transmembrane.

## Example 12 - lipoproteins

30

35

40

50

55

[0191] The incorporation of palmitate in recombinant lipoproteins was demonstrated by the method of Kraft et. al. [J. Bact. (1998) 180:3441-3447.]. Single colonies harbouring the plasmid of interest were grown overnight at 37°C in 20 ml of LB/Amp (100 $\mu$ g/ml) liquid culture. The culture was diluted to an OD<sub>550</sub> of 0.1 in 5.0 ml of fresh medium LB/Amp medium containing 5  $\mu$ C/ml [³H] palmitate (Amersham). When the OD<sub>550</sub> of the culture reached 0.4-0.8, recombinant lipoprotein was induced for 1 hour with IPTG (final concentration 1.0 mM). Bacteria were harvested by centrifugation in a bench top centrifuge at 2700g for 15 min and washed twice with 1.0 ml cold PBS. Cells were resuspended in 120 $\mu$ l of 20 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1.0% w/v SDS and lysed by boiling for 10 min. After centrifugation at 13000g for 10 min the supernatant was collected and proteins precipitated by the addition of 1.2 ml cold acetone and left for 1 hour at -20 °C. Protein was pelleted by centrifugation at 13000g for 10 min and resuspended in 20-50 $\mu$ l (calculated to standardise loading with respect to the final O.D of the culture) of 1.0% w/v SDS. An aliquot of 15  $\mu$ l was boiled with 5 $\mu$ l of SDS-PAGE sample buffer and analysed by SDS-PAGE. After electrophoresis gels were fixed for 1 hour in 10% v/v acetic acid and soaked for 30 minutes in Amplify solution (Amersham). The gel was vacuum-dried under heat and exposed to Hyperfilm (Kodak) overnight -80 °C.

[0192] Incorporation of the [<sup>3</sup>H] palmitate label, confirming lipidation, was found for the following proteins: Orf4L, Orf25L, 287L, 287LOrf4, 406.L, 576L, 926L, 919L and 919LOrf4.

#### Example 13 - domains in 287

**[0193]** Based on homology of different regions of 287 to proteins that belong to different functional classes, it was split into three 'domains', as shown in Figure 5. The second domain shows homology to IgA proteases, and the third domain shows homology to transferrin-binding proteins.

[0194] Each of the three 'domains' shows a different degree of sequence conservation between *N.meningitidis* strains - domain C is 98% identical, domain A is 83% identical, whilst domain B is only 71% identical. Note that protein 287 in strain MC58 is 61 amino acids longer than that of strain 2996. An alignment of the two sequences is shown in Figure

7, and alignments for various strains are disclosed in WO00/66741 (see Figures 5 and 15 therein).

[0195] The three domains were expressed individually as C-terminal His-tagged proteins. This was done for the MC58 and 2996 strains, using the following constructs:

287a-MC58 (aa 1-202), 287b-MC58 (aa 203-288), 287c-MC58 (aa 311-488).

287a-2996 (aa 1-139), 287b-2996 (aa 140-225), 287c-2996 (aa 250-427).

**[0196]** To make these constructs, the stop codon sequence was omitted in the 3'-end primer sequence. The 5' primers included the *Nhe*I restriction site, and the 3' primers included a *Xho*I as a tail, in order to direct the cloning of each amplified fragment into the expression vector pET21 b+ using *NdeI-Xho*I, *NheI-Xho*I or *NdeI-Hind*III restriction sites.

[0197] All six constructs could be expressed, but 287b-MC8 required denaturation and refolding for solubilisation.

[0198] Deletion of domain A is described below (' $\Delta 4$  287-His').

[0199] Immunological data (serum bactericidal assay) were also obtained using the various domains from strain 2996, against the homologous and heterologous MenB strains, as well as MenA (F6124 strain) and MenC (BZ133 strain):

	2996	BZ232	MC58	NGH38	394/98	MenA	MenC
287-His	32000	16	4096	4096	512	8000	16000
287(B)-His	256	-	-	-	-	16	-
287(C)-His	256	-	32	512	32	2048	>2048
287(B-C)-His	64000	128	4096	64000	1024	64000	32000

[0200] Using the domains of strain MC58, the following results were obtained:

	MC58	2996	BZ232	NGH38	394/98	MenA	MenC
287-His	4096	32000	16	4096	512	8000	16000
287(B)-His	128	128	-	-	-	-	128
287(C)-His	-	16	-	1024	-	512	-
287(B-C)-His	16000	64000	128	64000	512	64000	>8000

## Example 14 — deletions in 287

10

15

20

25

30

40

55

[0201] As well as expressing individual domains, 287 was also expressed (as a C-terminal His-tagged protein) by making progressive deletions within the first domain. These

[0202] Four deletion mutants of protein 287 from strain 2996 were used (Figure 6):

- 1) '287-His', consisting of amino acids 18-427 (i.e. leader peptide deleted);
- 2) '\(\Delta\)1 287-His', consisting of amino acids 26-427;
- 3)  $\Delta 2$  287-His, consisting of amino acids 70-427;
- 4) 'Δ3 287-His', consisting of amino acids 107-427; and
- 5) 'Δ4 287-His', consisting of amino acids 140-427 (=287-bc).
- <sup>45</sup> **[0203]** The 'Δ4' protein was also made for strain MC58 (' $\Delta$ 4 287MC58-His'; aa 203-488).

[0204] The constructs were made in the same way as 287a/b/c, as described above.

**[0205]** All six constructs could be expressed and protein could be purified. Expression of 287-His was, however, quite poor.

[0206] Expression was also high when the C-terminal His-tags were omitted.

[0207] Immunological data (serum bactericidal assay) were also obtained using the deletion mutants, against the homologous (2996) and heterologous MenB strains, as well as MenA (F6124 strain) and MenC (BZ133 strain):

	2996	BZ232	MC58	NGH38	394/98	MenA	MenC
287-his	32000	16	4096	4096	512	8000	16000
∆1 287-His	16000	128	4096	4096	1024	8000	16000
∆ <b>2 287-</b> His	16000	128	4096	>2048	512	16000	>8000

## (continued)

	2996	BZ232	MC58	NGH38	394/98	MenA	MenC
∆3 287-His	16000	128	4096	>2048	512	16000	>8000
∆4 287-His	64000	128	4096	64000	1024	64000	32000

[0208] The same high activity for the  $\Delta 4$  deletion was seen using the sequence from strain MC58.

[0209] As well as showing superior expression characteristics, therefore, the mutants are immunologically equivalent or superior.

## Example 15 - poly-glycine deletions

**[0210]** The ' $\Delta$ 1 287-His' construct of the previous example differs from 287-His and from '287<sup>untagged'</sup> only by a short N-terminal deletion (GGGGGGS). Using an expression vector which replaces the deleted serine with a codon present in the *Nhe* cloning site, however, this amounts to a deletion only of (Gly)<sub>6</sub>. Thus, the deletion of this (Gly)<sub>6</sub> sequence has been shown to have a dramatic effect on protein expression.

**[0211]** The protein lacking the N-terminal amino acids up to GGGGGG is called ' $\Delta$ G 287'. In strain MC58, its sequence (leader peptide underlined) is:

1	ΔG28	7	
0	סשעומם		

1	MFKRSVIAMA	CIFALSACGG	GGGGSPDVKS	ADTLSKPAAP	VVSEKETEAK
51	EDAPQAGSQG	QGAPSAQGSQ	DMAAVSEENT	GNGGAVTADN	PKNEDEVAQN
101	DMPQNAAGTD	SSTPNHTPDP	NMLAGNMENQ	ATDAGESSQP	ANQPDMANAA
151	DGMQGDDPSA	GGQNAGNTAA	QGANQAGNNQ	AAGSSDPIPA	SNPAPANGGS
201	NFGRVDLANG	VLIDGPSQNI	TLTHCKGDSC	SGNNFLDEEV	QLKSEFEKLS
251	DADKISNYKK	DGKNDKFVGL	VADSVQMKGI	NQYIIFYKPK	PTSFARFRRS
301	ARSRRSLPAE	MPLIPVNQAD	TLIVDGEAVS	LTGHSGNIFA	PEGNYRYLTY
351	GAEKLPGGSY	ALRVQGEPAK	GEMLAGAAVY	NGEVLHFHTE	NGRPYPTRGR
401	FAAKVDFGSK	SVDGIIDSGD	DLHMGTQKFK	AAIDGNGFKG	TWTENGSGDV
451	SGKFYGPAGE	EVAGKYSYRP	TDAEKGGFGV	FAGKKEOD*	

[0212] ΔG287, with or without His-tag ('ΔG287-His' and 'ΔG287K', respectively), are expressed at very good levels in comparison with the '287-His' or '287 untagged'.

[0213] On the basis of gene variability data, variants of  $\triangle$ G287-His were expressed in *E.coli* from a number of MenB strains, in particular from strains 2996, MC58, 1000, and BZ232. The results were also good.

**[0214]** It was hypothesised that poly-Gly deletion might be a general strategy to improve expression. Other MenB lipoproteins containing similar (Gly)<sub>n</sub> motifs (near the N-terminus, downstream of a cysteine) were therefore identified, namely Tbp2 (NMB0460), 741 (NMB 1870) and 983 (NMB1969):

40

45

50

5

10

20

25

30

35

TBP2			( ΔG1	bp2	
1	MNNPLVNQAA	MVLPVFLLSA	CLGGGGSFDL	DSVDTEAPRP	APKYQDVFSE
51	KPQAQKDQGG	YGFAMRLKRR	NWYPQAKEDE	VKLDESDWEA	TGLPDEPKEL
101	PKRQKSVIEK	VETDSDNNIY	SSPYLKPSNH	QNGNTGNGIN	QPKNQAKDYE
151	NFKYVYSGWF	YKHAKREFNL	KVEPKSAKNG	DDGYIFYHGK	EPSRQLPASG
201	KITYKGVWHF	ATDTKKGQKF	REIIQPSKSQ	GDRYSGFSGD	DGEEYSNKNK
251	STLTDGQEGY	GFTSNLEVDF	HNKKLTGKLI	RNNANTDNNQ	ATTTQYYSLE
301	AQVTGNRFNG	KATATDKPQQ	NSETKEHPFV	SDSSSLSGGF	FGPQGEELGF
351	RFLSDDQKVA	VVGSAKTKDK	PANGNTAAAS	GGTDAAASNG	AAGTSSENGK
401	LTTVLDAVEL	KLGDKEVQKL	DNFSNAAQLV	VDGIMIPLLP	EASESGNNQA

	451	NQGTNGGTAF TRKFDHTPES DKKDAQAGTQ TNGAQTASNT AGDTNGK	TKT
	501	YEVEVCCSNL NYLKYGMLTR KNSKSAMQAG ESSSQADAKT EQVEQSM	FLQ
	551	GERTDEKEIP SEQNIVYRGS WYGYIANDKS TSWSGNASNA TSGNRAE	
5	601	NFADKKITGT LTADNRQEAT FTIDGNIKDN GFEGTAKTAE SGFDLDQ	SNT
3	651	TRTPKAYITD AKVOGGFYGP KAEELGGWFA YPGDKQTKNA TNASGNS	
	701	VVFGAKROOP VR*	
10	741	( ΔG741	
	1	VNRTAFCCLS LTTALILTAC SSGGGGVAAD IGAGLADALT APLDHKDKO	GL
	51	QSLTLDQSVR KNEKLKLAAQ GAEKTYGNGD SLNTGKLKND KVSRFDFI	RQ
	101	IEVDGQLITL ESGEFQVYKQ SHSALTAFQT EQIQDSEHSG KMVAKRQF	RI
	151	GDIAGEHTSF DKLPEGGRAT YRGTAFGSDD AGGKLTYTID FAAKQGNG	ΚI
15	201	EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGG	KA
13	251	QEVAGSAEVK TVNGIRHIGL AAKQ*	
	983	( ΔG983	
20	1	MRTTPTFPTK TFKPTAMALA VATTLSACLG GGGGGTSAPD FNAGGTG	IGS
	51	NSRATTAKSA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAP	PPN
	101	LHTGDFPNPN DAYKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSIS	
	151	LYGRKEHGYN ENYKNYTAYM RKEAPEDGGG KDIEASFDDE AVIETEA	
	201	DIRHVKEIGH IDLVSHIIGG RSVDGRPAGG IAPDATLHIM NTNDETK	
25	251	MVAAIRNAWV KLGERGVRIV NNSFGTTSRA GTADLFQIAN SEEQYRQA	
23	301	DYSGGDKTDE GIRLMQQSDY GNLSYHIRNK NMLFIFSTGN DAQAQPN	
	351	LLPFYEKDAQ KGIITVAGVD RSGEKFKREM YGEPGTEPLE YGSNHCG	
	401	MWCLSAPYEA SVRFTRTNPI QIAGTSFSAP IVTGTAALLL QKYPWMS	
	451	LRTTLLTTAQ DIGAVGVDSK FGWGLLDAGK AMNGPASFPF GDFTADTI	
	501	SDIAYSFRND ISGTGGLIKK GGSQLQLHGN NTYTGKTIIE GGSLVLY	
30	551	KSDMRVETKG ALIYNGAASG GSLNSDGIVY LADTDQSGAN ETVHIKG	_
	601	LDGKGTLYTR LGKLLKVDGT AIIGGKLYMS ARGKGAGYLN STGRRVP	
	651	AAKIGQDYSF FTNIETDGGL LASLDSVEKT AGSEGDTLSY YVRRGNAM	
	701	ASAAAHSAPA GLKHAVEQGG SNLENLMVEL DASESSATPE TVETAAAI	
	751	DMPGIRPYGA TFRAAAAVQH ANAADGVRIF NSLAATVYAD STAAHADN	
	801	RRLKAVSDGL DHNGTGLRVI AQTQQDGGTW EQGGVEGKMR GSTQTVG	
35	851	KTGENTTAAA TLGMGRSTWS ENSANAKTDS ISLFAGIRHD AGDIGYL	
	901	FSYGRYKNSI SRSTGADEHA EGSVNGTLMQ LGALGGVNVP FAATGDL	
	951	GGLRYDLLKQ DAFAEKGSAL GWSGNSLTEG TLVGLAGLKL SQPLSDKA	AVL

1051 FGNGWNGLAR YSYAGSKQYG NHSGRVGVGY RF\*

40

45

50

55

**[0215]** Tbp2 and 741 genes were from strain MC58; 983 and 287 genes were from strain 2996. These were cloned in pET vector and expressed in *E.coli* without the sequence coding for their leader peptides or as " $\Delta$ G forms", both fused to a C-terminal His-tag. In each case, the same effect was seen - expression was good in the clones carrying the deletion of the poly-glycine stretch, and poor or absent if the glycines were present in the expressed protein:

1001 FATAGVERDL NGRDYTVTGG FTGATAATGK TGARNMPHTR LVAGLGADVE

ORF	Express.	Purification	Bact. Activity
287-His(2996)	+/-	+	+
'287untagged'(2996)	+/-	nd	nd
∆G287-His(2996)	+	+	+
∆G287K(2996)	+	+	+
∆G287-His(MC58)	+	+	+
∆G287-His(1000)	+	+	+
∆G287-His(B <b>Z</b> 232)	+	+	+
Tbp2-His(MC58)	+/-	nd	nd
∆GTbp2-His(MC58)	+	+	
741-His(MC58)	+/-	nd	nd

(continued)

ORF	Express.	Purification	Bact. Activity
∆G741-His(MC58)	+	+	
983-His (2996)			
∆G983-His (2996)	+	+	

[0216] SDS-PAGE of the proteins is shown in Figure 13.

 $\Delta G287$  and hybrids

5

10

40

45

50

55

**[0217]**  $\Delta$ G287 proteins were made and purified for strains MC58, 1000 and BZ232. Each of these gave high ELISA titres and also serum bactericidal titres of>8192.  $\Delta$ G287K, expressed from pET-24b, gave excellent titres in ELISA and the serum bactericidal assay.  $\Delta$ G287-ORF46.1K may also be expressed in pET-24b.

 $\begin{tabular}{ll} \textbf{[0218]} & \Delta G287 \ was also fused directly in-frame upstream of 919, 953, 961 \end{tabular} \label{eq:continuous} (sequences shown below) and ORF46.1: \begin{tabular}{ll} \textbf{(Sequences shown below)} & \textbf{(Sequences shown bel$ 

	ΔG287-9	<u>19</u>				
	1	ATGGCTAGCC	CCGATGTTAA	ATCGGCGGAC	ACGCTGTCAA	AACCGGCCGC
20	51	TCCTGTTGTT	GCTGAAAAAG	AGACAGAGGT	AAAAGAAGAT	GCGCCACAGG
	101	CAGGTTCTCA	AGGACAGGGC	GCGCCATCCA	CACAAGGCAG	CCAAGATATG
	151	GCGGCAGTTT	CGGCAGAAAA	TACAGGCAAT	GGCGGTGCGG	CAACAACGGA
	201	CAAACCCAAA	AATGAAGACG	AGGGACCGCA	AAATGATATG	CCGCAAAATT
	251	CCGCCGAATC	CGCAAATCAA	ACAGGGAACA	ACCAACCCGC	CGATTCTTCA
	301	GATTCCGCCC	CCGCGTCAAA	CCCTGCACCT	GCGAATGGCG	GTAGCAATTT
25	351	TGGAAGGGTT	GATTTGGCTA	ATGGCGTTTT	GATTGATGGG	CCGTCGCAAA
	401	ATATAACGTT	GACCCACTGT	AAAGGCGATT	CTTGTAATGG	TGATAATTTA
	451	TTGGATGAAG	AAGCACCGTC	AAAATCAGAA	TTTGAAAATT	TAAATGAGTC
	501	TGAACGAATT	GAGAAATATA	AGAAAGATGG	GAAAAGCGAT	AAATTTACTA
	551	ATTTGGTTGC	GACAGCAGTT	CAAGCTAATG	GAACTAACAA	ATATGTCATC
30	601	ATTTATAAAG	ACAAGTCCGC	TTCATCTTCA	TCTGCGCGAT	TCAGGCGTTC
30	651	TGCACGGTCG	AGGAGGTCGC	TTCCTGCCGA	GATGCCGCTA	ATCCCCGTCA
	701	ATCAGGCGGA	TACGCTGATT	GTCGATGGGG	AAGCGGTCAG	CCTGACGGGG
	751	CATTCCGGCA	ATATCTTCGC	GCCCGAAGGG	AATTACCGGT	ATCTGACTTA
	801	CGGGGCGGAA	AAATTGCCCG	GCGGATCGTA	TGCCCTCCGT	GTGCAAGGCG
	851	AACCGGCAAA	AGGCGAAATG	CTTGCTGGCA	CGGCCGTGTA	CAACGGCGAA
35	901	GTGCTGCATT	TTCATACGGA	AAACGGCCGT	CCGTACCCGA	CTAGAGGCAG
	951	GTTTGCCGCA	AAAGTCGATT	TCGGCAGCAA	ATCTGTGGAC	GGCATTATCG
	1001	ACAGCGGCGA	TGATTTGCAT	ATGGGTACGC	AAAAATTCAA	AGCCGCCATC

	1051	GATGGAAACG	GCTTTAAGGG	GACTTGGACG	GAAAATGGCG	GCGGGGATGT
	1101	TTCCGGAAGG	TTTTACGGCC	CGGCCGGCGA	GGAAGTGGCG	GGAAAATACA
	1151	GCTATCGCCC	GACAGATGCG	GAAAAGGGCG	GATTCGGCGT	GTTTGCCGGC
5	1201	AAAAAAGAGC	AGGATGGATC	CGGAGGAGGA	GGATGCCAAA	GCAAGAGCAT
3	1251	CCAAACCTTT	CCGCAACCCG	ACACATCCGT	CATCAACGGC	CCGGACCGGC
	1301	CGGTCGGCAT	CCCCGACCCC	GCCGGAACGA	CGGTCGGCGG	CGGCGGGGCC
	1351	GTCTATACCG	TTGTACCGCA	CCTGTCCCTG	CCCCACTGGG	CGGCGCAGGA
	1401	TTTCGCCAAA	AGCCTGCAAT	CCTTCCGCCT	CGGCTGCGCC	AATTTGAAAA
	1451	ACCGCCAAGG	CTGGCAGGAT	GTGTGCGCCC	AAGCCTTTCA	AACCCCCGTC
10	1501	CATTCCTTTC	AGGCAAAACA	GTTTTTTGAA	CGCTATTTCA	CGCCGTGGCA
	1551	GGTTGCAGGC	AACGGAAGCC	TTGCCGGTAC	GGTTACCGGC	TATTACGAGC
	1601	CGGTGCTGAA	GGGCGACGAC	AGGCGGACGG	CACAAGCCCG	CTTCCCGATT
	1651	TACGGTATTC	CCGACGATTT	TATCTCCGTC	CCCCTGCCTG	CCGGTTTGCG
	1701		GCCCTTGTCC			
	1751	CAATCGACAA	TACCGGCGGC	ACACATACCG	CCGACCTCTC	CCGATTCCCC
15	1801	ATCACCGCGC	GCACAACGGC	AATCAAAGGC	AGGTTTGAAG	GAAGCCGCTT
	1851	CCTCCCCTAC	CACACGCGCA	ACCAAATCAA	CGGCGGCGCG	CTTGACGGCA
	1901	AAGCCCCGAT	ACTCGGTTAC	GCCGAAGACC	CCGTCGAACT	TTTTTTTATG
	1951	CACATCCAAG	GCTCGGGCCG	TCTGAAAACC	CCGTCCGGCA	AATACATCCG
	2001	CATCGGCTAT	GCCGACAAAA	ACGAACATCC	CTACGTTTCC	ATCGGACGCT
00	2051	ATATGGCGGA	CAAAGGCTAC	CTCAAGCTCG	GGCAGACCTC	GATGCAGGGC
20	2101	ATCAAAGCCT	ATATGCGGCA	AAATCCGCAA	CGCCTCGCCG	AAGTTTTGGG
	2151	TCAAAACCCC	AGCTATATCT	TTTTCCGCGA	GCTTGCCGGA	AGCAGCAATG
	2201	ACGGTCCCGT	CGGCGCACTG	GGCACGCCGT	TGATGGGGGA	ATATGCCGGC
	2251	GCAGTCGACC	GGCACTACAT	TACCTTGGGC	GCGCCCTTAT	TTGTCGCCAC
	2301		GTTACCCGCA			
25	2351	ATACCGGCAG	CGCGATTAAA	GGCGCGGTGC	GCGTGGATTA	TTTTTGGGGA
	2401	TACGGCGACG	AAGCCGGCGA	ACTTGCCGGC	AAACAGAAAA	CCACGGGTTA
	2451	CGTCTGGCAG	CTCCTACCCA	ACGGTATGAA	GCCCGAATAC	CGCCCGTAAC
	2501	TCGAG				
30						
	1		TLSKPAAPVV			
	51		GGAATTDKPK	_	-	_
	101		ANGGSNFGRV			
	151		FENLNESERI			
35	201		SARFRRSARS		_	
	251		NYRYLTYGAE PYPTRGRFAA		-	
	301 351		ENGGGDVSGR			
	401		GCQSKSIQTF			
	451		PHWAAQDFAK			
	501		RYFTPWOVAG	_		
40	551		PLPAGLRSGK			-
	601		RFEGSRFLPY			
	651		PSGKYIRIGY			
	701		RLAEVLGQNP			
	751		APLFVATAHP			
45	801		KOKTTGYVWO			
,,,						

	ΔG287-9	53				
	1		CCGATGTTAA	ATCGGCGGAC	ACGCTGTCAA	AACCGGCCGC
	51		GCTGAAAAAG			
5	101		AGGACAGGGC			
	151		CGGCAGAAAA			
	201		AATGAAGACG			
	251		CGCAAATCAA			
	301		CCGCGTCAAA			
	351	TGGAAGGGTT	GATTTGGCTA	ATGGCGTTTT	GATTGATGGG	CCGTCGCAAA
10	401	ATATAACGTT	GACCCACTGT	AAAGGCGATT	CTTGTAATGG	TGATAATTTA
	451	TTGGATGAAG	AAGCACCGTC	AAAATCAGAA	TTTGAAAATT	TAAATGAGTC
	501	TGAACGAATT	GAGAAATATA	AGAAAGATGG	GAAAAGCGAT	AAATTTACTA
	551	ATTTGGTTGC	GACAGCAGTT	CAAGCTAATG	GAACTAACAA	ATATGTCATC
	601	ATTTATAAAG	ACAAGTCCGC	TTCATCTTCA	TCTGCGCGAT	TCAGGCGTTC
15	651	TGCACGGTCG	AGGAGGTCGC	TTCCTGCCGA	GATGCCGCTA	ATCCCCGTCA
15	701	ATCAGGCGGA	TACGCTGATT	GTCGATGGGG	AAGCGGTCAG	CCTGACGGGG
	751	CATTCCGGCA	ATATCTTCGC	GCCCGAAGGG	AATTACCGGT	ATCTGACTTA
20						
	801	CGGGGCGGAA	AAATTGCCCG	GCGGATCGTA	TGCCCTCCGT	GTGCAAGGCG
	851		AGGCGAAATG			
	901		TTCATACGGA			
	951		AAAGTCGATT			
	1001	ACAGCGGCGA	TGATTTGCAT	ATGGGTACGC	AAAAATTCAA	AGCCGCCATC
25	1051		GCTTTAAGGG			
	1101	TTCCGGAAGG	TTTTACGGCC	CGGCCGGCGA	GGAAGTGGCG	GGAAAATACA
	1151	GCTATCGCCC	GACAGATGCG	GAAAAGGGCG	GATTCGGCGT	GTTTGCCGGC
	1201	AAAAAAGAGC	AGGATGGATC	CGGAGGAGGA	GGAGCCACCT	ACAAAGTGGA
	1251		GCCAACGCCC			
30	1301	CCAACGTCGG	CGGTTTTTAC	GGTCTGACCG	GTTCCGTCGA	GTTCGACCAA
30	1351	GCAAAACGCG	ACGGTAAAAT	CGACATCACC	ATCCCCGTTG	CCAACCTGCA
	1401	AAGCGGTTCG	CAACACTTTA	CCGACCACCT	GAAATCAGCC	GACATCTTCG
	1451	ATGCCGCCCA	ATATCCGGAC	ATCCGCTTTG	TTTCCACCAA	ATTCAACTTC
	1501		AACTGGTTTC			
	1551	AACCGCCCCC	GTCAAACTCA	AAGCCGAAAA	ATTCAACTGC	TACCAAAGCC
35	1601		AACCGAAGTT			
	1651		GGGGCGTGGA			
	1701	CGTCCGCATC	GACATCCAAA	TCGAGGCAGC	CAAACAATAA	CTCGAG
40	1	MASPDVKSAD	TLSKPAAPVV	AEKETEVKED	APQAGSQGQG	APSTQGSQDM
	51		GGAATTDKPK			
	101	DSAPASNPAP	ANGGSNFGRV	DLANGVLIDG	PSQNITLTHC	KGDSCNGDNL
	151		FENLNESERI			
	201	IYKDKSASSS	SARFRRSARS	RRSLPAEMPL	IPVNQADTLI	VDGEAVSLTG
45	251		NYRYLTYGAE			
45	301		PYPTRGRFAA			_
	351		ENGGGDVSGR			
	401	_	GATYKVDEYH			
	451		IPVANLQSGS			
	501		NLTMHGKTAP		YQSPMAKTEV	CGGDFSTTID
50	551	KTKWGVDYLV	NVGMTKSVRI	DIQIEAAKQ*		

	ΔG287-9	61				
	1		CCGATGTTAA	ATCGGCGGAC	ACGCTGTCAA	AACCGGCCGC
	51			AGACAGAGGT		
	101	CAGGTTCTCA	AGGACAGGGC	GCGCCATCCA	CACAAGGCAG	CCAAGATATG
5	151	GCGGCAGTTT	CGGCAGAAAA	TACAGGCAAT	GGCGGTGCGG	CAACAACGGA
	201	CAAACCCAAA	AATGAAGACG	AGGGACCGCA	AAATGATATG	CCGCAAAATT
	251	CCGCCGAATC	CGCAAATCAA	ACAGGGAACA	ACCAACCCGC	CGATTCTTCA
	301	GATTCCGCCC	CCGCGTCAAA	CCCTGCACCT	GCGAATGGCG	GTAGCAATTT
	351	TGGAAGGGTT	GATTTGGCTA	ATGGCGTTTT	GATTGATGGG	CCGTCGCAAA
10	401	ATATAACGTT	GACCCACTGT	AAAGGCGATT	CTTGTAATGG	TGATAATTTA
70	451	TTGGATGAAG	AAGCACCGTC	AAAATCAGAA	TTTGAAAATT	TAAATGAGTC
	501	TGAACGAATT	GAGAAATATA	AGAAAGATGG	GAAAAGCGAT	AAATTTACTA
	551	ATTTGGTTGC	GACAGCAGTT	CAAGCTAATG	GAACTAACAA	ATATGTCATC
	601	ATTTATAAAG	ACAAGTCCGC	TTCATCTTCA	TCTGCGCGAT	TCAGGCGTTC
	651	TGCACGGTCG	AGGAGGTCGC	TTCCTGCCGA	GATGCCGCTA	ATCCCCGTCA
1 <i>5</i>	701	ATCAGGCGGA	TACGCTGATT	GTCGATGGGG	AAGCGGTCAG	CCTGACGGGG
	751	CATTCCGGCA	ATATCTTCGC	GCCCGAAGGG	AATTACCGGT	ATCTGACTTA
	801	CGGGGCGGAA	AAATTGCCCG	GCGGATCGTA	TGCCCTCCGT	GTGCAAGGCG
	851	AACCGGCAAA	AGGCGAAATG	CTTGCTGGCA	CGGCCGTGTA	CAACGGCGAA
	901	GTGCTGCATT	TTCATACGGA	AAACGGCCGT	CCGTACCCGA	CTAGAGGCAG
	951			TCGGCAGCAA		
20	1001			ATGGGTACGC		
	1051			GACTTGGACG		
	1101			CGGCCGGCGA		
	1151			GAAAAGGGCG		
	1201			CGGAGGAGGA		
25	1251			TGGCCATTGC		
23	1301			GCTGGAGAGA		
	1351			AGACGCAACT		
	1401			TGAAAAAAGT		
	1451			AACGTCGATG		
	1501			AACCAAGTTA		
30	1551	AGCAGATACT	GATGCCGCTC	TGGATGCAAC	CACCAACGCC	TTGAATAAAT
	1601					AAATATCGTA
35	1651			AGCCGTGGCT		
	1701			CCGATTCATT		
	1751			GCCAATGAAG		
	1801			CAAAGTAAAA		
	1851			GCACAGCTAA		
40	1901			ACCGACATCA		
40	1951			AGCAAACAGT		
	2001					GCTACTACCG
	2051					TGCCGATCAC
	2101					TGCGCAAAGA
	2151					CTGTTCCAAC
45	2201					CGGCTACAAA
	2251					CCGAAAACTT
	2301			TACGAGTGGT		TCTTCCGCAG
	2351	CCIACCAIGT	CGGCGTCAAT	TACGAGIGGT	AACICGAG	

	1	MASPDVKSAD	TLSKPAAPVV	AEKETEVKED	APQAGSQGQG	APSTQGSQDM
	51	AAVSAENTGN	GGAATTDKPK	NEDEGPQNDM	PQNSAESANQ	TGNNQPADSS
	101	DSAPASNPAP	ANGGSNFGRV	DLANGVLIDG	PSQNITLTHC	KGDSCNGDNL
	151	LDEEAPSKSE	FENLNESERI	EKYKKDGKSD	KFTNLVATAV	QANGTNKYVI
5	201	IYKDKSASSS	SARFRRSARS	RRSLPAEMPL	IPVNQADTLI	VDGEAVSLTG
	251	HSGNIFAPEG	NYRYLTYGAE	KLPGGSYALR	VQGEPAKGEM	LAGTAVYNGE
	301	VLHFHTENGR	PYPTRGRFAA	KVDFGSKSVD	GIIDSGDDLH	MGTQKFKAAI
	351	DGNGFKGTWT	ENGGGDVSGR	FYGPAGEEVA	GKYSYRPTDA	EKGGFGVFAG
	401	KKEQDGSGGG	GATNDDDVKK	AATVAIAAAY	NNGQEINGFK	AGETIYDIDE
10	451	DGTITKKDAT	AADVEADDFK	GLGLKKVVTN	LTKTVNENKQ	NVDAKVKAAE
10	501	SEIEKLTTKL	ADTDAALADT	DAALDATTNA	LNKLGENITT	FAEETKTNIV
	551	KIDEKLEAVA	DTVDKHAEAF	NDIADSLDET	NTKADEAVKT	ANEAKQTAEE
	601	TKQNVDAKVK	AAETAAGKAE	AAAGTANTAA	DKAEAVAAKV	TDIKADIATN
	651	KDNIAKKANS	ADVYTREESD	SKFVRIDGLN	ATTEKLDTRL	ASAEKSIADH
	701	DTRLNGLDKT	VSDLRKETRQ	GLAEQAALSG	LFQPYNVGRF	NVTAAVGGYK
15	751	SESAVAIGTG	FRFTENFAAK	AGVAVGTSSG	SSAAYHVGVN	YEW*

	ELISA	Bactericidal
ΔG287-953-His	3834	65536
∆G287-961-His	108627	65536

[0219] The bactericidal efficacy (homologous strain) of antibodies raised against the hybrid proteins was compared with antibodies raised against simple mixtures of the component antigens (using 287-GST) for 919 and ORF46.1:

	Mixture with 287	Hybrid with ∆G287
919	32000	128000
ORF46.1	128	16000

[0220] Data for bactericidal activity against heterologous MenB strains and against serotypes A and C were also obtained:

	91	9	ORF46.1	
Strain	Mixture	Hybrid	Mixture	Hybrid
NGH38	1024	32000	-	16384
MC58	512	8192	-	512
BZ232	512	512	-	-
MenA (F6124)	512	32000	-	8192
MenC (C11)	>2048	>2048	-	-
MenC (BZ133)	>4096	64000	-	8192

**[0221]** The hybrid proteins with  $\Delta$ G287 at the N-terminus are therefore immunologically superior to simple mixtures, with  $\Delta$ G287-ORF46.1 being particularly effective, even against heterologous strains.  $\Delta$ G287-ORF46.1K may be expressed in pET-24b.

[0222] The same hybrid proteins were made using New Zealand strain 394/98 rather than 2996:

	4.000.7Ng	01.0				
	ΔG287NZ	<del></del>	CCGATGTCAA	GTCGGCGGAC	ΔCCCTCTCΔΔ	AACCTGCCGC
	51		TCTGAAAAAG			
	101		AGGACAGGGC			
5	151		CGGAAGAAAA			
	201	CAAACCCAAA	AATGAAGACG	AGGGGGCGCA	AAATGATATG	CCGCAAAATG
	251	CCGCCGATAC	AGATAGTTTG	ACACCGAATC	ACACCCCGGC	TTCGAATATG
	301		ATATGGAAAA			
	351	GCCGGCAAAC	CAACCGGATA	TGGCAAATAC	GGCGGACGGA	ATGCAGGGTG
10	401		GGCAGGCGGG			
10	451		CCGAAAACAA			
	501		CCTAGCGCCA			
	551		TTCTGTTGTG			
	601		AAGGCGATTC			
	651		AAATCAGAAT			
15	701		GAAAGATGGG			
	751 801		CCGATAGTGT CCTAAACCCA			
	851		GTCGCTTCCG			
	901		TGATTGTCGA			
	951		TTCGCGCCCG			
20	1001		GCCCGGCGGA			
20	1051		AAATGCTCGC			
	1101		ACGGAAAACG			
	1151		CGATTTCGGC			
	1201	GGCGATGGTT	TGCATATGGG	TACGCAAAAA	TTCAAAGCCG	CCATCGATGG
	1251	AAACGGCTTT	AAGGGGACTT	GGACGGAAAA	TGGCGGCGGG	GATGTTTCCG
25	1301		CGGCCCGGCC			
	1351	CGCCCAACAG	ATGCGGAAAA	GGGCGGATTC	GGCGTGTTTG	CCGGCAAAAA
	1401		GGATCCGGAG			
	1451		ACCCGACACA			
	1501		ACCCCGCCGG			
30	1551		CCGCACCTGT			
	1601		GCAATCCTTC			
	1651		AGGATGTGTG			
	1701 1751		AAACAGTTTT AAGCCTTGCC			
	1801		ACGACAGGCG			
	1851		GATTTTATCT			
35	1901		TGTCCGCATC			
	1951		GCGGCACACA			
	2001		ACGGCAATCA			
		00000001011	11000001211011	12.0000011		
40						
	2051	CCTACCACAC	GCGCAACCAA	ATCAACGGCG	GCGCGCTTGA	CGGCAAAGCC
	2101		GTTACGCCGA			
	2151		GGCCGTCTGA			
	2201		CAAAAACGAA			
	2251		GCTACCTCAA			
45	2301		CGGCAAAATC			
	2351		TATCTTTTTC			
	2401		CACTGGGCAC			
	2451		TACATTACCT			
	2501		CCGCAAAGCC			
50	2551		TTAAAGGCGC			
50	2601		GGCGAACTTG			
	2651		ACCCAACGGT			

	1	MASPDVKSAD	TLSKPAAPVV	SEKETEAKED	APOAGSOGOG	APSAOGGODM
	51		GGAAATDKPK			
	101	PAGNMENQAP	DAGESEQPAN	QPDMANTADG	MQGDDPSAGG	ENAGNTAAQG
_	151	TNQAENNQTA	GSQNPASSTN	PSATNSGGDF	GRTNVGNSVV	IDGPSQNITL
5	201	THCKGDSCSG	NNFLDEEVQL	KSEFEKLSDA	DKISNYKKDG	KNDGKNDKFV
	251	GLVADSVQMK	GINQYIIFYK	PKPTSFARFR	RSARSRRSLP	AEMPLIPVNQ
	301	ADTLIVDGEA	VSLTGHSGNI	FAPEGNYRYL	TYGAEKLPGG	SYALRVQGEP
	351		VYNGEVLHFH			
	401	GDGLHMGTQK	FKAAIDGNGF	KGTWTENGGG	DVSGKFYGPA	GEEVAGKYSY
10	451	RPTDAEKGGF	GVFAGKKEQD	GSGGGGCQSK	SIQTFPQPDT	SVINGPDRPV
70	501	GIPDPAGTTV	GGGGAVYTVV	PHLSLPHWAA	QDFAKSLQSF	RLGCANLKNR
	551	QGWQDVCAQA	FQTPVHSFQA	KQFFERYFTP	WQVAGNGSLA	GTVTGYYEPV
	601	LKGDDRRTAQ	ARFPIYGIPD	DFISVPLPAG	LRSGKALVRI	RQTGKNSGTI
	651	DNTGGTHTAD	LSRFPITART	TAIKGRFEGS	RFLPYHTRNQ	INGGALDGKA
	701	PILGYAEDPV	ELFFMHIQGS	GRLKTPSGKY	IRIGYADKNE	HPYVSIGRYM
15	751	ADKGYLKLGQ	TSMQGIKAYM	RQNPQRLAEV	LGQNPSYIFF	RELAGSSNDG
	801	PVGALGTPLM	GEYAGAVDRH	YITLGAPLFV	ATAHPVTRKA	LNRLIMAQDT
	851	GSAIKGAVRV	DYFWGYGDEA	GELAGKQKTT	GYVWQLLPNG	MKPEYRP*
20						
	ΔG287N2	2-953				
	1	ATGGCTAGCC	CCGATGTCAA	GTCGGCGGAC	ACGCTGTCAA	AACCTGCCGC
	51	CCCTGTTGTT	TCTGAAAAAG	AGACAGAGGC	AAAGGAAGAT	GCGCCACAGG
	101	CAGGTTCTCA	AGGACAGGGC	GCGCCATCCG	CACAAGGCGG	TCAAGATATG
0.5	151		CGGAAGAAAA			
25	201	CAAACCCAAA	AATGAAGACG	AGGGGGCGCA	AAATGATATG	CCGCAAAATG
	251	CCGCCGATAC	AGATAGTTTG	ACACCGAATC	ACACCCCGGC	TTCGAATATG
	301	CCGGCCGGAA	ATATGGAAAA	CCAAGCACCG	GATGCCGGGG	AATCGGAGCA
	351	GCCGGCAAAC	CAACCGGATA	TGGCAAATAC	GGCGGACGGA	ATGCAGGGTG
	401	ACGATCCGTC	GGCAGGCGGG	GAAAATGCCG	GCAATACGGC	TGCCCAAGGT
30	451		CCGAAAACAA			
00	501	TTCAACCAAT	CCTAGCGCCA	CGAATAGCGG	TGGTGATTTT	GGAAGGACGA
	551	ACGTGGGCAA	TTCTGTTGTG	ATTGACGGGC	CGTCGCAAAA	TATAACGTTG
	601	ACCCACTGTA	AAGGCGATTC	TTGTAGTGGC	AATAATTTCT	TGGATGAAGA
	651	AGTACAGCTA	AAATCAGAAT	TTGAAAAATT	AAGTGATGCA	GACAAAATAA
	701	GTAATTACAA	GAAAGATGGG	AAGAATGACG	GGAAGAATGA	TAAATTTGTC
35	751	GGTTTGGTTG	CCGATAGTGT	GCAGATGAAG	GGAATCAATC	AATATATTAT
	801		CCTAAACCCA			
	851	GGTCGAGGCG	GTCGCTTCCG	GCCGAGATGC	CGCTGATTCC	CGTCAATCAG
	901	GCGGATACGC	TGATTGTCGA	TGGGGAAGCG	GTCAGCCTGA	CGGGGCATTC
	951	CGGCAATATC	TTCGCGCCCG	AAGGGAATTA	CCGGTATCTG	ACTTACGGGG
	1001	CGGAAAAATT	GCCCGGCGGA	TCGTATGCCC	TCCGTGTTCA	AGGCGAACCT
40	1051	TCAAAAGGCG	AAATGCTCGC	GGGCACGGCA	GTGTACAACG	GCGAAGTGCT
	1101	GCATTTTCAT	ACGGAAAACG	GCCGTCCGTC	CCCGTCCAGA	GGCAGGTTTG
	1151	CCGCAAAAGT	CGATTTCGGC	AGCAAATCTG	TGGACGGCAT	TATCGACAGC
	1201	GGCGATGGTT	TGCATATGGG	TACGCAAAAA	TTCAAAGCCG	CCATCGATGG
	1251	AAACGGCTTT	AAGGGGACTT	GGACGGAAAA	TGGCGGCGGG	GATGTTTCCG
.=	1301		CGGCCCGGCC			
45	1351		ATGCGGAAAA			
	1401		GGATCCGGAG			
	1451		CGCCCGTTTC			
	1501		TTTACGGTCT			
	1551		AAAATCGACA			
	= 2 3 =					

5	1601 1651 1701 1751 1801 1851	GCCCAATATC CAAAAAACTG CCCCCGTCAA GCGAAAACCG CAAATGGGGC	CGGACATCCG GTTTCCGTTG ACTCAAAGCC AAGTTTGCGG GTGGACTACC	CACCTGAAAT CTTTGTTTCC ACGGCAACCT GAAAAATTCA CGGCGACTTC TCGTTAACGT GCAGCCAAAC	ACCAAATTCA GACCATGCAC ACTGCTACCA AGCACCACCA TGGTATGACC	ACTTCAACGG GGCAAAACCG AAGCCCGATG TCGACCGCAC AAAAGCGTCC
10	1 51 101	AAVSEENTGN	GGAAATDKPK	SEKETEAKED NEDEGAQNDM QPDMANTADG	PQNAADTDSL	TPNHTPASNM
15	151 201 251 301 351 401	TNQAENNQTA THCKGDSCSG GLVADSVQMK ADTLIVDGEA SKGEMLAGTA	GSQNPASSTN NNFLDEEVQL GINQYIIFYK VSLTGHSGNI VYNGEVLHFH	PSATNSGGDF KSEFEKLSDA PKPTSFARFR FAPEGNYRYL TENGRPSPSR KGTWTENGGG	GRTNVGNSVV DKISNYKKDG RSARSRRSLP TYGAEKLPGG GRFAAKVDFG	IDGPSQNITL KNDGKNDKFV AEMPLIPVNQ SYALRVQGEP SKSVDGIIDS
20	451 501 551 601	VGGFYGLTGS AQYPDIRFVS	VEFDQAKRDG TKFNFNGKKL	GSGGGGATYK KIDITIPVAN VSVDGNLTMH VDYLVNVGMT	LQSGSQHFTD GKTAPVKLKA	HLKSADIFDA EKFNCYQSPM
25						
30						
35						
40						
45						
50						
55						

	ΔG287NZ-	-961				
	1		CCGATGTCAA	GTCGGCGGAC	ACGCTGTCAA	AACCTGCCGC
	51	CCCTGTTGTT	TCTGAAAAAG	AGACAGAGGC	AAAGGAAGAT	GCGCCACAGG
5	101	CAGGTTCTCA	AGGACAGGGC	GCGCCATCCG	CACAAGGCGG	TCAAGATATG
	151	GCGGCGGTTT	CGGAAGAAAA	TACAGGCAAT	GGCGGTGCGG	CAGCAACGGA
	201			AGGGGGCGCA		
	251	CCGCCGATAC	AGATAGTTTG	ACACCGAATC	ACACCCGGC	TTCGAATATG
	301	CCGGCCGGAA	ATATGGAAAA	CCAAGCACCG	GATGCCGGGG	AATCGGAGCA
	351	GCCGGCAAAC	CAACCGGATA	TGGCAAATAC	GGCGGACGGA	ATGCAGGGTG
10	401	ACGATCCGTC	GGCAGGCGGG	GAAAATGCCG	GCAATACGGC	TGCCCAAGGT
	451	ACAAATCAAG	CCGAAAACAA	TCAAACCGCC	GGTTCTCAAA	ATCCTGCCTC
	501	TTCAACCAAT	CCTAGCGCCA	CGAATAGCGG	TGGTGATTTT	GGAAGGACGA
	551	ACGTGGGCAA	TTCTGTTGTG	ATTGACGGGC	CGTCGCAAAA	TATAACGTTG
	601	ACCCACTGTA	AAGGCGATTC	TTGTAGTGGC	AATAATTTCT	TGGATGAAGA
15	651	AGTACAGCTA	AAATCAGAAT	TTGAAAAATT	AAGTGATGCA	GACAAAATAA
13	701	GTAATTACAA	GAAAGATGGG	AAGAATGACG	GGAAGAATGA	TAAATTTGTC
	751	GGTTTGGTTG	CCGATAGTGT	GCAGATGAAG	GGAATCAATC	AATATATTAT
	801	CTTTTATAAA	CCTAAACCCA	CTTCATTTGC	GCGATTTAGG	CGTTCTGCAC
	851	GGTCGAGGCG	GTCGCTTCCG	GCCGAGATGC	CGCTGATTCC	CGTCAATCAG
	901	GCGGATACGC	TGATTGTCGA	TGGGGAAGCG	GTCAGCCTGA	CGGGGCATTC
20	951	CGGCAATATC	TTCGCGCCCG	AAGGGAATTA	CCGGTATCTG	ACTTACGGGG
	1001			TCGTATGCCC		
	1051	TCAAAAGGCG	AAATGCTCGC	GGGCACGGCA	GTGTACAACG	GCGAAGTGCT
	1101	GCATTTTCAT	ACGGAAAACG	GCCGTCCGTC	CCCGTCCAGA	GGCAGGTTTG
	1151	CCGCAAAAGT	CGATTTCGGC	AGCAAATCTG	TGGACGGCAT	TATCGACAGC
	1201			TACGCAAAAA		
25	1251			GGACGGAAAA		
	1301	GAAAGTTTTA	CGGCCCGGCC	GGCGAGGAAG	TGGCGGGAAA	ATACAGCTAT
	1351	CGCCCAACAG	ATGCGGAAAA	GGGCGGATTC	GGCGTGTTTG	CCGGCAAAAA
	1401	AGAGCAGGAT	GGATCCGGAG	GAGGAGGAGC	CACAAACGAC	GACGATGTTA
	1451	AAAAAGCTGC	CACTGTGGCC	ATTGCTGCTG	CCTACAACAA	TGGCCAAGAA
	1501	ATCAACGGTT	TCAAAGCTGG	AGAGACCATC	TACGACATTG	ATGAAGACGG
30	1551	CACAATTACC	AAAAAAGACG	CAACTGCAGC	CGATGTTGAA	GCCGACGACT
	1601			AAAGTCGTGA		
	1651			CGATGCCAAA		
	1701			AGTTAGCAGA		
	1751			GCAACCACCA		
35	1801			TGAAGAGACT		
	1851			TGGCTGATAC		
	1901			TCATTGGATG		
	1951			TGAAGCCAAA		
	2001			TAAAAGCTGC		
	2051			GCTAATACTG		
40	2101	GTCGCTGCAA	AAGTTACCGA	CATCAAAGCT	GATATCGCTA	CGAACAAAGA
	2151	TAATATTGCT	AAAAAAGCAA	ACAGTGCCGA	CGTGTACACC	AGAGAAGAGT
45	2201	CTGACAGCAA	ATTTGTCAGA	ATTGATGGTC	TGAACGCTAC	TACCGAAAAA
	2251	TTGGACACAC	GCTTGGCTTC	TGCTGAAAAA	TCCATTGCCG	ATCACGATAC
	2301	TCGCCTGAAC	GGTTTGGATA	AAACAGTGTC	AGACCTGCGC	AAAGAAACCC
	2351	GCCAAGGCCT	TGCAGAACAA	GCCGCGCTCT	CCGGTCTGTT	CCAACCTTAC
	2401	AACGTGGGTC	GGTTCAATGT	AACGGCTGCA	GTCGGCGGCT	ACAAATCCGA
50	2451	ATCGGCAGTC	GCCATCGGTA	CCGGCTTCCG	CTTTACCGAA	AACTTTGCCG
50	2501	CCAAAGCAGG	CGTGGCAGTC	GGCACTTCGT	CCGGTTCTTC	CGCAGCCTAC
	2551	CATGTCGGCG	TCAATTACGA	GTGGTAAAAG	CTT	

	1	MASPDVKSAD	TLSKPAAPVV	SEKETEAKED	APQAGSQGQG	APSAQGGQDM
	51	AAVSEENTGN	GGAAATDKPK	NEDEGAQNDM	PQNAADTDSL	TPNHTPASNM
	101	PAGNMENQAP	DAGESEQPAN	QPDMANTADG	MQGDDPSAGG	ENAGNTAAQG
5	151	TNQAENNQTA	GSQNPASSTN	PSATNSGGDF	GRTNVGNSVV	IDGPSQNITL
	201	THCKGDSCSG	NNFLDEEVQL	KSEFEKLSDA	DKISNYKKDG	KNDGKNDKFV
	251	GLVADSVQMK	GINQYIIFYK	PKPTSFARFR	RSARSRRSLP	AEMPLIPVNQ
	301	ADTLIVDGEA	VSLTGHSGNI	FAPEGNYRYL	TYGAEKLPGG	SYALRVQGEP
	351	${\tt SKGEMLAGTA}$	VYNGEVLHFH	TENGRPSPSR	GRFAAKVDFG	SKSVDGIIDS
	401	GDGLHMGTQK	FKAAIDGNGF	KGTWTENGGG	DVSGKFYGPA	GEEVAGKYSY
10	451	RPTDAEKGGF	GVFAGKKEQD	GSGGGGATND	DDVKKAATVA	IAAAYNNGQE
	501	INGFKAGETI	YDIDEDGTIT	KKDATAADVE	ADDFKGLGLK	KVVTNLTKTV
	551	NENKQNVDAK	VKAAESEIEK	LTTKLADTDA	ALADTDAALD	ATTNALNKLG
	601	ENITTFAEET	KTNIVKIDEK	LEAVADTVDK	HAEAFNDIAD	SLDETNTKAD
	651	EAVKTANEAK	QTAEETKQNV	DAKVKAAETA	AGKAEAAAGT	ANTAADKAEA
15	701	VAAKVTDIKA	DIATNKDNIA	KKANSADVYT	REESDSKFVR	IDGLNATTEK
15	751	LDTRLASAEK	SIADHDTRLN	GLDKTVSDLR	KETRQGLAEQ	AALSGLFQPY
	801	NVGRFNVTAA	VGGYKSESAV	AIGTGFRFTE	NFAAKAGVAV	GTSSGSSAAY
	851	HVGVNYEW*				

# $^{20}$ $\Delta G983$ and hybrids

25

30

55

[0223] Bactericidal titres generated in response to  $\Delta G983$  (His-fusion) were measured against various strains, including the homologous 2996 strain:

	2996	NGH38	BZ133
∆G983	512	128	128

[0224]  $\Delta$ G983 was also expressed as a hybrid, with ORF46.1, 741, 961 or 961c at its C-terminus:

	ΔG983-01	RF46.1				
	1	ATGACTTCTG	CGCCCGACTT	CAATGCAGGC	GGTACCGGTA	TCGGCAGCAA
	51	CAGCAGAGCA	ACAACAGCGA	AATCAGCAGC	AGTATCTTAC	GCCGGTATCA
35	101	AGAACGAAAT	GTGCAAAGAC	AGAAGCATGC	TCTGTGCCGG	TCGGGATGAC
00	151	GTTGCGGTTA	CAGACAGGGA	TGCCAAAATC	AATGCCCCCC	CCCCGAATCT
	201	GCATACCGGA	GACTTTCCAA	ACCCAAATGA	CGCATACAAG	AATTTGATCA
	251	ACCTCAAACC	TGCAATTGAA	GCAGGCTATA	CAGGACGCGG	GGTAGAGGTA
	301	GGTATCGTCG	ACACAGGCGA	ATCCGTCGGC	AGCATATCCT	TTCCCGAACT
	351	GTATGGCAGA	AAAGAACACG	GCTATAACGA	AAATTACAAA	AACTATACGG
40	401	CGTATATGCG	GAAGGAAGCG	CCTGAAGACG	GAGGCGGTAA	AGACATTGAA
	451	GCTTCTTTCG	ACGATGAGGC	CGTTATAGAG	ACTGAAGCAA	AGCCGACGGA
	501	TATCCGCCAC	GTAAAAGAAA	TCGGACACAT	CGATTTGGTC	TCCCATATTA
	551	TTGGCGGGCG	TTCCGTGGAC	GGCAGACCTG	CAGGCGGTAT	TGCGCCCGAT
	601	GCGACGCTAC	ACATAATGAA	TACGAATGAT	GAAACCAAGA	ACGAAATGAT
	651	GGTTGCAGCC	ATCCGCAATG	CATGGGTCAA	GCTGGGCGAA	CGTGGCGTGC
45	701	GCATCGTCAA	TAACAGTTTT	GGAACAACAT	CGAGGGCAGG	CACTGCCGAC
	751	CTTTTCCAAA	TAGCCAATTC	GGAGGAGCAG	TACCGCCAAG	CGTTGCTCGA
	801	CTATTCCGGC	GGTGATAAAA	CAGACGAGGG	TATCCGCCTG	ATGCAACAGA
	851	GCGATTACGG	CAACCTGTCC	TACCACATCC	GTAATAAAAA	CATGCTTTTC
50	901	ATCTTTTCGA	CAGGCAATGA	CGCACAAGCT	CAGCCCAACA	CATATGCCCT
	951	ATTGCCATTT	TATGAAAAAG	ACGCTCAAAA	AGGCATTATC	ACAGTCGCAG
	1001	GCGTAGACCG	CAGTGGAGAA	AAGTTCAAAC	GGGAAATGTA	TGGAGAACCG
	1051	GGTACAGAAC	CGCTTGAGTA	TGGCTCCAAC	CATTGCGGAA	TTACTGCCAT
	1101	GTGGTGCCTG	TCGGCACCCT	ATGAAGCAAG	CGTCCGTTTC	ACCCGTACAA

	1151			ACATCCTTTT		
	1201	ACGGCGGCTC	TGCTGCTGCA	GAAATACCCG	TGGATGAGCA	ACGACAACCT
	1251	GCGTACCACG	TTGCTGACGA	CGGCTCAGGA	CATCGGTGCA	GTCGGCGTGG
	1301	ACAGCAAGTT	CGGCTGGGGA	CTGCTGGATG	CGGGTAAGGC	CATGAACGGA
5	1351	CCCGCGTCCT	TTCCGTTCGG	CGACTTTACC	GCCGATACGA	AAGGTACATC
	1401	CGATATTGCC	TACTCCTTCC	GTAACGACAT	TTCAGGCACG	GGCGGCCTGA
	1451			CTGCAACTGC		
	1501			CGGTTCGCTG		
	1551			CCAAAGGTGC		
	1601			AGCGACGGCA		
10	1651			AACCGTACAC		
	1701	GGACGGCAAA	GGTACGCTGT	ACACACGTTT	GGGCAAACTG	CTGAAAGTGG
	1751	ACGGTACGGC	GATTATCGGC	GGCAAGCTGT	ACATGTCGGC	ACGCGGCAAG
	1801	GGGGCAGGCT	ATCTCAACAG	TACCGGACGA	CGTGTTCCCT	TCCTGAGTGC
	1851	CGCCAAAATC	GGGCAGGATT	ATTCTTTCTT	CACAAACATC	GAAACCGACG
	1901			GACAGCGTCG		
45	1951			TGTCCGTCGC		
15						
	2001			CGCCCGCCGG		
	2051			AACCTGATGG		
	2101			GGTTGAAACT		
	2151			ACGGCGCAAC		
	2201	TACAGCATGC	GAATGCCGCC	GACGGTGTAC	GCATCTTCAA	CAGTCTCGCC
20	2251	GCTACCGTCT	ATGCCGACAG	TACCGCCGCC	CATGCCGATA	TGCAGGGACG
20	2301	CCGCCTGAAA	GCCGTATCGG	ACGGGTTGGA	CCACAACGGC	ACGGGTCTGC
	2351	GCGTCATCGC	GCAAACCCAA	CAGGACGGTG	GAACGTGGGA	ACAGGGCGGT
	2401			CAGTACCCAA		
	2451			CAGCCGCCAC		
	2501			AATGCAAAAA		
	2551			GGGCGATATC		
25				ACAGCATCAG		
	2601					
	2651			AACGGCACGC		
	2701			TGCCGCAACG		
	2751			TCAAACAGGA		
	2801	GTGCTTTGGG	CTGGAGCGGC	AACAGCCTCA	CTGAAGGCAC	GCTGGTCGGA
	2851	CTCGCGGGTC	TGAAGCTGTC	GCAACCCTTG	AGCGATAAAG	CCGTCCTGTT
30	2901	TGCAACGGCG	GGCGTGGAAC	GCGACCTGAA	CGGACGCGAC	TACACGGTAA
	2951	CGGGCGGCTT	TACCGGCGCG	ACTGCAGCAA	CCGGCAAGAC	GGGGGCACGC
	3001	AATATGCCGC	ACACCCGTCT	GGTTGCCGGC	CTGGGCGCGG	ATGTCGAATT
	3051	CGGCAACGGC	TGGAACGGCT	TGGCACGTTA	CAGCTACGCC	GGTTCCAAAC
	3101			CGAGTCGGCG		
	3151			CTCAGATTTG		
35	3201			ATTTCGAACC		
				GCCGAGCGCA		
	3251					
	3301			GGGCAACCTG		
	3351			TTGTCCGCTT		
	3401			CATGCCTCAC		
	3451			TAGCCTTTAC		
40	3501	CGAACACCAT	CCCGCCGACG	GCTATGACGG	GCCACAGGGC	GGCGGCTATC
	3551	CCGCTCCCAA	AGGCGCGAGG	GATATATACA	GCTACGACAT	AAAAGGCGTT
	3601			CCTGACCGAC		
	3651	GCTTGCCGAC	CGTTTCCACA	ATGCCGGTAG	TATGCTGACG	CAAGGAGTAG
	3701			ACCCGATACA		
	3751			CAACGGCACT		
45	3801			TTGTCGGCGC		
45	3851			GCTGTCATGC		
	3901			CATCAACGAT		
	3951			CCATCCGCGA		
	4001			GCCGTCAGCA		
	4051			TGTTCGGGGA		
50	4101	CACGGCACAT	CCTATCAAGC	GGTCGCAGAT	GGGCGCGATC	GCATTGCCGA
	4151	AAGGGAAATC	CGCCGTCAGC	GACAATTTTG	CCGATGCGGC	ATACGCCAAA
	4201	TACCCGTCCC	CTTACCATTC	CCGAAATATC	CGTTCAAACT	TGGAGCAGCG
	4251			CCTCCTCAAC		
	4301			CAACGCCACC		
	4351			TTTTGAGAAG		
	4401		CACCACCACC		J.10010111111	
55	1401	CONGCHUCHU	CACCACCACC	AUTUA		

	1	MTSADDENAC	GTGTGSNSDA	TTAKSAAVSY	<b>VCIKNEWCKD</b>	R SMT.C ACRDD
	51			DEPNPNDAYK		
	101			KEHGYNENYK		
	151			VKEIGHIDLV		
5	201			IRNAWVKLGE		
	251			GDKTDEGIRL		
	301		-	YEKDAOKGII		
	351			SAPYEASVRF		
	401			LLTTAODIGA		
4.0	451			YSFRNDISGT		
10	501			RVETKGALIY		
	551			GTLYTRLGKL		
	601			GODYSFFTNI		
	651	GDTLSYYVRR	GNAARTASAA	AHSAPAGLKH	AVEQGGSNLE	NLMVELDASE
	701	SSATPETVET	AAADRTDMPG	IRPYGATFRA	AAAVQHANAA	DGVRIFNSLA
15	751	ATVYADSTAA	HADMQGRRLK	AVSDGLDHNG	TGLRVIAQTQ	QDGGTWEQGG
	801	VEGKMRGSTQ	TVGIAAKTGE	NTTAAATLGM	GRSTWSENSA	NAKTDSISLF
	851	AGIRHDAGDI	GYLKGLFSYG	RYKNSISRST	GADEHAEGSV	NGTLMQLGAL
	901	GGVNVPFAAT	GDLTVEGGLR	YDLLKQDAFA	EKGSALGWSG	NSLTEGTLVG
	951	LAGLKLSQPL	SDKAVLFATA	GVERDLNGRD	YTVTGGFTGA	TAATGKTGAR
	1001	NMPHTRLVAG	LGADVEFGNG	WNGLARYSYA	GSKQYGNHSG	RVGVGYRFLD
20	1051	GGGGTGSSDL	ANDSFIRQVL	DRQHFEPDGK	YHLFGSRGEL	AERSGHIGLG
	1101	KIQSHQLGNL	MIQQAAIKGN	IGYIVRFSDH	GHEVHSPFDN	HASHSDSDEA
	1151	GSPVDGFSLY	RIHWDGYEHH	PADGYDGPQG	GGYPAPKGAR	DIYSYDIKGV
	1201	AQNIRLNLTD	NRSTGQRLAD	RFHNAGSMLT	QGVGDGFKRA	TRYSPELDRS
	1251			AGEIVGAGDA	-	
25	1301		_	AAAAIRDWAV		
23	1351			PIKRSQMGAI		
	1401			ENITSSTVPP	SNGKNVKLAD	QRHPKTGVPF
	1451	DGKGFPNFEK	HVKYDTLEHH	нннн*		

		41				
_	1	ATGACTTCTG	CGCCCGACTT	CAATGCAGGC	GGTACCGGTA	TCGGCAGCAA
	51	CAGCAGAGCA	ACAACAGCGA	AATCAGCAGC	AGTATCTTAC	GCCGGTATCA
5	101	AGAACGAAAT	GTGCAAAGAC	AGAAGCATGC	TCTGTGCCGG	TCGGGATGAC
	151	GTTGCGGTTA	CAGACAGGGA	TGCCAAAATC	AATGCCCCCC	CCCCGAATCT
	201	GCATACCGGA	GACTTTCCAA	ACCCAAATGA	CGCATACAAG	AATTTGATCA
	251	ACCTCAAACC	TGCAATTGAA	GCAGGCTATA	CAGGACGCGG	GGTAGAGGTA
	301	GGTATCGTCG	ACACAGGCGA	ATCCGTCGGC	AGCATATCCT	TTCCCGAACT
	351	GTATGGCAGA	AAAGAACACG	GCTATAACGA	AAATTACAAA	AACTATACGG
10	401	CGTATATGCG	GAAGGAAGCG	CCTGAAGACG	GAGGCGGTAA	AGACATTGAA
	451	GCTTCTTTCG	ACGATGAGGC	CGTTATAGAG	ACTGAAGCAA	AGCCGACGGA
	501	TATCCGCCAC	GTAAAAGAAA	TCGGACACAT	CGATTTGGTC	TCCCATATTA
	551	TTGGCGGGCG	TTCCGTGGAC	GGCAGACCTG	CAGGCGGTAT	TGCGCCCGAT
	601	GCGACGCTAC	ACATAATGAA	TACGAATGAT	GAAACCAAGA	ACGAAATGAT
	651	GGTTGCAGCC	ATCCGCAATG	CATGGGTCAA	GCTGGGCGAA	CGTGGCGTGC
15	701	GCATCGTCAA	TAACAGTTTT	GGAACAACAT	CGAGGGCAGG	CACTGCCGAC
	751	CTTTTCCAAA	TAGCCAATTC	GGAGGAGCAG	TACCGCCAAG	CGTTGCTCGA
	801	CTATTCCGGC	GGTGATAAAA	CAGACGAGGG	TATCCGCCTG	ATGCAACAGA
	851	GCGATTACGG	CAACCTGTCC	TACCACATCC	GTAATAAAAA	CATGCTTTTC
	901	ATCTTTTCGA	CAGGCAATGA	CGCACAAGCT	CAGCCCAACA	CATATGCCCT
20	951	ATTGCCATTT	TATGAAAAAG	ACGCTCAAAA	AGGCATTATC	ACAGTCGCAG
20	1001	GCGTAGACCG	CAGTGGAGAA	AAGTTCAAAC	GGGAAATGTA	TGGAGAACCG
	1051	GGTACAGAAC	CGCTTGAGTA	TGGCTCCAAC	CATTGCGGAA	TTACTGCCAT
	1101	GTGGTGCCTG	TCGGCACCCT	ATGAAGCAAG	CGTCCGTTTC	ACCCGTACAA
	1151	ACCCGATTCA	AATTGCCGGA	ACATCCTTTT	CCGCACCCAT	CGTAACCGGC
	1201	ACGGCGGCTC	${\tt TGCTGCTGCA}$	GAAATACCCG	TGGATGAGCA	ACGACAACCT
25	1251	GCGTACCACG	TTGCTGACGA	CGGCTCAGGA	CATCGGTGCA	GTCGGCGTGG
	1301	ACAGCAAGTT	CGGCTGGGGA	CTGCTGGATG	CGGGTAAGGC	CATGAACGGA
	1351	CCCGCGTCCT	TTCCGTTCGG	CGACTTTACC	GCCGATACGA	AAGGTACATC
	1401	CGATATTGCC	TACTCCTTCC	GTAACGACAT	TTCAGGCACG	GGCGGCCTGA
	1451	TCAAAAAAGG	CGGCAGCCAA	CTGCAACTGC	ACGGCAACAA	CACCTATACG
	1501	GGCAAAACCA	TTATCGAAGG	CGGTTCGCTG	GTGTTGTACG	GCAACAACAA
30	1551	ATCGGATATG	CGCGTCGAAA	CCAAAGGTGC	GCTGATTTAT	AACGGGGCGG
	1601	CATCCGGCGG	CAGCCTGAAC	AGCGACGGCA	TTGTCTATCT	GGCAGATACC
	1651	GACCAATCCG	GCGCAAACGA	AACCGTACAC	ATCAAAGGCA	GTCTGCAGCT

	1701	GGACGGCAAA	GGTACGCTGT	ACACACGTTT	GGGCAAACTG	CTGAAAGTGG
	1751	ACGGTACGGC	GATTATCGGC	GGCAAGCTGT	ACATGTCGGC	ACGCGGCAAG
	1801		ATCTCAACAG			
5	1851		GGGCAGGATT			
	1901		GGCTTCCCTC			
	1951	GGCGACACGC	TGTCCTATTA	TGTCCGTCGC	GGCAATGCGG	CACGGACTGC
	2001	TTCGGCAGCG	GCACATTCCG	CGCCCGCCGG	TCTGAAACAC	GCCGTAGAAC
	2051	AGGGCGGCAG	CAATCTGGAA	AACCTGATGG	TCGAACTGGA	TGCCTCCGAA
	2101	TCATCCGCAA	CACCCGAGAC	GGTTGAAACT	GCGGCAGCCG	ACCGCACAGA
10	2151		ATCCGCCCCT			
	2201	TACAGCATGC	GAATGCCGCC	GACGGTGTAC	GCATCTTCAA	CAGTCTCGCC
	2251	GCTACCGTCT	ATGCCGACAG	TACCGCCGCC	CATGCCGATA	TGCAGGGACG
	2301	CCGCCTGAAA	GCCGTATCGG	ACGGGTTGGA	CCACAACGGC	ACGGGTCTGC
	2351	GCGTCATCGC	GCAAACCCAA	CAGGACGGTG	GAACGTGGGA	ACAGGGCGGT
4.5	2401	GTTGAAGGCA	AAATGCGCGG	CAGTACCCAA	ACCGTCGGCA	TTGCCGCGAA
15	2451	AACCGGCGAA	AATACGACAG	CAGCCGCCAC	ACTGGGCATG	GGACGCAGCA
	2501	CATGGAGCGA	AAACAGTGCA	AATGCAAAAA	CCGACAGCAT	TAGTCTGTTT
	2551	GCAGGCATAC	GGCACGATGC	GGGCGATATC	GGCTATCTCA	AAGGCCTGTT
	2601	CTCCTACGGA	CGCTACAAAA	ACAGCATCAG	CCGCAGCACC	GGTGCGGACG
	2651	AACATGCGGA	AGGCAGCGTC	AACGGCACGC	TGATGCAGCT	GGGCGCACTG
20	2701	GGCGGTGTCA	ACGTTCCGTT	TGCCGCAACG	GGAGATTTGA	CGGTCGAAGG
20	2751		TACGACCTGC			
	2801	GTGCTTTGGG	CTGGAGCGGC	AACAGCCTCA	CTGAAGGCAC	GCTGGTCGGA
	2851		TGAAGCTGTC			
	2901	TGCAACGGCG	GGCGTGGAAC	GCGACCTGAA	CGGACGCGAC	TACACGGTAA
	2951		TACCGGCGCG			
25	3001	AATATGCCGC	ACACCCGTCT	GGTTGCCGGC	CTGGGCGCGG	ATGTCGAATT
	3051	CGGCAACGGC	TGGAACGGCT	TGGCACGTTA	CAGCTACGCC	GGTTCCAAAC
	3101	AGTACGGCAA	CCACAGCGGA	CGAGTCGGCG	TAGGCTACCG	GTTCCTCGAG
	3151		GGGGTGGTGT			
	3201		GCACCGCTCG			
	3251		GTCCGTCAGG			
30	3301		AAACTTATGG			
	3351		AAGGTCAGCC			
	3401		CATTACCTTG			
	3451		CCTTAACCGC			
	3501		AAGATGGTTG			
	3551		TACATCTTTT			
35	3601		CGGCGTTCGG			
	3651		TTCGCCGCCA			
	3701		ACTCAATGTC			
	3751		ATGCCGTCAT			
	3801		TACTCCCTCG			
40	3851		GGAAGTGAAA			TATCGGCCTT
70	3901	GCCGCCAAGC	AACTCGAGCA	CCACCACCAC	CACCACTGA	

5 .	1 51 101 151 201 251 301	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ	GTGIGSNSRA NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF
10	351 401 451 501 551 601 651	TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR	HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA	LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI	VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL	LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE
15	701 751 801 851 901 951	SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL	AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA	IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD	AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA	DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR
20	1001	NMPHTRLVAG	LGADVEFGNG	WNGLARYSYA	GSKQYGNHSG	RVGVGYRFLE
25	1051 1101 1151 1201 1251 1301	GAEKTYGNGD SHSALTAFQT YRGTAFGSDD	IGAGLADALT SLNTGKLKND EQIQDSEHSG AGGKLTYTID VLYNQAEKGS HH*	KVSRFDFIRQ KMVAKRQFRI FAAKQGNGKI	IEVDGQLITL GDIAGEHTSF EHLKSPELNV	ESGEFQVYKQ DKLPEGGRAT DLAAADIKPD
30						
35						
40						
45						
50						
55						

	ΔG983-96	£1				
	1		CGCCCGACTT	CAATGCAGGC	GGTACCGGTA	TCGGCAGCAA
	51			AATCAGCAGC		
5	101			AGAAGCATGC		
	151			TGCCAAAATC		
	201			ACCCAAATGA		
	251	ACCTCAAACC	TGCAATTGAA	GCAGGCTATA	CAGGACGCGG	GGTAGAGGTA
	301	GGTATCGTCG	ACACAGGCGA	ATCCGTCGGC	AGCATATCCT	TTCCCGAACT
4.0	351	GTATGGCAGA	AAAGAACACG	GCTATAACGA	AAATTACAAA	AACTATACGG
10	401			CCTGAAGACG		
	451	GCTTCTTTCG	ACGATGAGGC	CGTTATAGAG	ACTGAAGCAA	AGCCGACGGA
	501	TATCCGCCAC	GTAAAAGAAA	TCGGACACAT	CGATTTGGTC	TCCCATATTA
	551	TTGGCGGGCG	TTCCGTGGAC	GGCAGACCTG	CAGGCGGTAT	TGCGCCCGAT
	601	GCGACGCTAC	ACATAATGAA	TACGAATGAT	GAAACCAAGA	ACGAAATGAT
15	651	GGTTGCAGCC	ATCCGCAATG	CATGGGTCAA	GCTGGGCGAA	CGTGGCGTGC
	701	GCATCGTCAA	TAACAGTTTT	GGAACAACAT	CGAGGGCAGG	CACTGCCGAC
	751	CTTTTCCAAA	TAGCCAATTC	GGAGGAGCAG	TACCGCCAAG	CGTTGCTCGA
	801	CTATTCCGGC	GGTGATAAAA	CAGACGAGGG	TATCCGCCTG	ATGCAACAGA
	851	GCGATTACGG	CAACCTGTCC	TACCACATCC	${\tt GTAATAAAAA}$	CATGCTTTTC
	901	ATCTTTTCGA	CAGGCAATGA	CGCACAAGCT	CAGCCCAACA	CATATGCCCT
20	951			ACGCTCAAAA		
	1001	GCGTAGACCG	CAGTGGAGAA	AAGTTCAAAC	GGGAAATGTA	TGGAGAACCG
	1051			TGGCTCCAAC		
	1101			ATGAAGCAAG		
	1151			ACATCCTTTT		
	1201			GAAATACCCG		
25	1251			CGGCTCAGGA		
	1301			CTGCTGGATG		
	1351			CGACTTTACC		
	1401			GTAACGACAT		
	1451			CTGCAACTGC		
30	1501			CGGTTCGCTG		
	1551			CCAAAGGTGC		
	1601			AGCGACGGCA		
	1651 1701			AACCGTACAC ACACACGTTT		
	1751			GGCAAGCTGT		
	1801			TACCGGACGA		
35	1851			ATTCTTTCTT		
	1901			GACAGCGTCG		
	1951			TGTCCGTCGC		
	2001			CGCCCGCCGG		
	2051			AACCTGATGG		
10	2101			GGTTGAAACT		
40	2151	TATGCCGGGC	ATCCGCCCCT	ACGGCGCAAC	TTTCCGCGCA	GCGGCAGCCG
	2201	TACAGCATGC	GAATGCCGCC	GACGGTGTAC	GCATCTTCAA	CAGTCTCGCC
	2251	GCTACCGTCT	ATGCCGACAG	TACCGCCGCC	CATGCCGATA	TGCAGGGACG
	2301			ACGGGTTGGA		
	2351	GCGTCATCGC	GCAAACCCAA	CAGGACGGTG	GAACGTGGGA	ACAGGGCGGT
45	2401	GTTGAAGGCA	AAATGCGCGG	CAGTACCCAA	ACCGTCGGCA	TTGCCGCGAA
	2451	AACCGGCGAA	AATACGACAG	CAGCCGCCAC	ACTGGGCATG	GGACGCAGCA
	2501	CATGGAGCGA	AAACAGTGCA	AATGCAAAAA	CCGACAGCAT	TAGTCTGTTT
	2551	GCAGGCATAC	GGCACGATGC	GGGCGATATC	GGCTATCTCA	AAGGCCTGTT
	2601	CTCCTACGGA	CGCTACAAAA	ACAGCATCAG	CCGCAGCACC	GGTGCGGACG
	2651			AACGGCACGC		
50	2701	GGCGGTGTCA	ACGTTCCGTT	TGCCGCAACG	GGAGATTTGA	CGGTCGAAGG
	2751	CGGTCTGCGC	TACGACCTGC	TCAAACAGGA	TGCATTCGCC	GAAAAAGGCA
	2801	GTGCTTTGGG	CTGGAGCGGC	AACAGCCTCA	CTGAAGGCAC	GCTGGTCGGA
	2851	CTCGCGGGTC	TGAAGCTGTC	GCAACCCTTG	AGCGATAAAG	CCGTCCTGTT

	2901	TGCAACGGCG	GGCGTGGAAC	GCGACCTGAA	CGGACGCGAC	TACACGGTAA
	2951		TACCGGCGCG			
	3001	AATATGCCGC	ACACCCGTCT	GGTTGCCGGC	CTGGGCGCGG	ATGTCGAATT
5	3051	CGGCAACGGC	TGGAACGGCT	TGGCACGTTA	CAGCTACGCC	GGTTCCAAAC
_	3101	AGTACGGCAA	CCACAGCGGA	CGAGTCGGCG	TAGGCTACCG	GTTCCTCGAG
	3151	GGTGGCGGAG	GCACTGGATC	CGCCACAAAC	GACGACGATG	TTAAAAAAGC
	3201	TGCCACTGTG	GCCATTGCTG	CTGCCTACAA	CAATGGCCAA	GAAATCAACG
	3251	GTTTCAAAGC	TGGAGAGACC	ATCTACGACA	TTGATGAAGA	CGGCACAATT
	3301	ACCAAAAAAG	ACGCAACTGC	AGCCGATGTT	GAAGCCGACG	ACTTTAAAGG
10	3351	TCTGGGTCTG	AAAAAAGTCG	TGACTAACCT	GACCAAAACC	GTCAATGAAA
	3401	ACAAACAAAA	CGTCGATGCC	AAAGTAAAAG	CTGCAGAATC	TGAAATAGAA
	3451	AAGTTAACAA	CCAAGTTAGC	AGACACTGAT	GCCGCTTTAG	CAGATACTGA
	3501	TGCCGCTCTG	GATGCAACCA	CCAACGCCTT	GAATAAATTG	GGAGAAAATA
	3551	TAACGACATT	TGCTGAAGAG	ACTAAGACAA	ATATCGTAAA	AATTGATGAA
	3601	AAATTAGAAG	CCGTGGCTGA	TACCGTCGAC	AAGCATGCCG	AAGCATTCAA
15	3651	CGATATCGCC	GATTCATTGG	ATGAAACCAA	CACTAAGGCA	GACGAAGCCG
	3701	TCAAAACCGC	CAATGAAGCC	AAACAGACGG	CCGAAGAAAC	CAAACAAAAC
	3751	GTCGATGCCA	AAGTAAAAGC	TGCAGAAACT	GCAGCAGGCA	AAGCCGAAGC
	3801	TGCCGCTGGC	ACAGCTAATA	CTGCAGCCGA	CAAGGCCGAA	GCTGTCGCTG
	3851	CAAAAGTTAC	CGACATCAAA	GCTGATATCG	CTACGAACAA	AGATAATATT
	3901	GCTAAAAAAG	CAAACAGTGC	CGACGTGTAC	ACCAGAGAAG	AGTCTGACAG
20	3951	CAAATTTGTC	AGAATTGATG	GTCTGAACGC	TACTACCGAA	AAATTGGACA
	4001	CACGCTTGGC	TTCTGCTGAA	AAATCCATTG	CCGATCACGA	TACTCGCCTG
	4051	AACGGTTTGG	ATAAAACAGT	GTCAGACCTG	CGCAAAGAAA	CCCGCCAAGG
	4101	CCTTGCAGAA	CAAGCCGCGC	TCTCCGGTCT	GTTCCAACCT	TACAACGTGG
	4151	GTCGGTTCAA	TGTAACGGCT	GCAGTCGGCG	GCTACAAATC	CGAATCGGCA
25	4201	GTCGCCATCG	GTACCGGCTT	CCGCTTTACC	GAAAACTTTG	CCGCCAAAGC
23	4251		GTCGGCACTT			
	4301	GCGTCAATTA	CGAGTGGCTC	GAGCACCACC	ACCACCACCA	CTGA
30	1		GTGIGSNSRA			
30	1 51	VAVTDRDAKI	NAPPPNLHTG	DFPNPNDAYK	NLINLKPAIE	AGYTGRGVEV
30	51 101	VAVTDRDAKI GIVDTGESVG	NAPPPNLHTG SISFPELYGR	DFPNPNDAYK KEHGYNENYK	NLINLKPAIE NYTAYMRKEA	AGYTGRGVEV PEDGGGKDIE
30	51 101 151	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE	NAPPPNLHTG SISFPELYGR TEAKPTDIRH	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD
30	51 101 151 201	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD
	51 101 151 201 251	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF
<i>30</i>	51 101 151 201 251 301	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP
	51 101 151 201 251	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG
	51 101 151 201 251 301 351 401	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG
	51 101 151 201 251 301 351 401 451	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT
	51 101 151 201 251 301 351 401 451 501	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT
35	51 101 151 201 251 301 351 401 451 501	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK
	51 101 151 201 251 301 351 401 451 501 551 601	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE
35	51 101 151 201 251 301 351 401 451 501 551 601 651	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE
35	51 101 151 201 251 301 351 401 451 501 551 601 651 701	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA
35	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG
35	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF
35 40	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL
35	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG
35 40	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR
35 40	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL NMPHTRLVAG	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA LGADVEFGNG	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD WNGLARYSYA	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA GSKQYGNHSG	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR RVGVGYRFLE
35 40	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951 1001	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL NMPHTRLVAG GGGGTGSATN	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA LGADVEFGNG DDDVKKAATV	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD WNGLARYSYA AIAAAYNNGQ	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA GSKQYGNHSG EINGFKAGET	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR RVGVGYRFLE IYDIDEDGTI
35 40	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951 1001 1051	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL NMPHTRLVAG GGGGTGSATN TKKDATAADV	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA LGADVEFGNG DDDVKKAATV EADDFKGLGL	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD WNGLARYSYA AIAAAYNNGQ KKVVTNLTKT	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA GSKQYGNHSG EINGFKAGET VNENKQNVDA	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LUDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR RVGVGYRFLE IYDIDEDGTI KVKAAESEIE
35 40	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951 1001 1051 1101	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL NMPHTRLVAG GGGGTGSATN TKKDATAADV KLTTKLADTD	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA LGADVEFGNG DDDVKKAATV EADDFKGLGL AALADTDAAL	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD WNGLARYSYA AIAAAYNNGQ KKVVTNLTKT DATTNALNKL	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA GSKQYGNHSG EINGFKAGET VNENKQNVDA GENITTFAEE	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR RVGVGYRFLE IYDIDEDGTI KVKAAESEIE TKTNIVKIDE
35 40 45	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951 1001 1051 1101 1151 1201	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL NMPHTRLVAG GGGGTGSATN TKKDATAADV KLTTKLADTD KLEAVADTVD	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA LGADVEFGNG DDDVKKAATV EADDFKGLGL AALADTDAAL KHAEAFNDIA	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD WNGLARYSYA AIAAAYNNGQ KKVVTNLTKT DATTNALNKL DSLDETNTKA	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA GSKQYGNHSG EINGFKAGET VNENKQNVDA GENITTFAEE DEAVKTANEA	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR RVGVGYRFLE IYDIDEDGTI KVKAAESEIE TKTNIVKIDE KQTAEETKQN
35 40 45	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951 1001 1051 1101 1151 1201 1251	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL NMPHTRLVAG GGGGTGSATN TKKDATAADV KLTTKLADTD KLEAVADTVD VDAKVKAAET	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA LGADVEFGNG DDDVKKAATV EADDFKGLGL AALADTDAAL KHAEAFNDIA AAGKAEAAAG	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD WNGLARYSYA AIAAAYNNGQ KKVVTNLTKT DATTNALNKL DSLDETNTKA TANTAADKAE	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA GSKQYGNHSG EINGFKAGET VNENKQNVDA GENITTFAEE DEAVKTANEA AVAAKVTDIK	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR RVGVGYRFLE IYDIDEDGTI KVKAAESEIE TKTNIVKIDE KQTAEETKQN ADIATNKDNI
35 40 45	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951 1001 1051 1101 1151 1201 1251 1301	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL NMPHTRLVAG GGGGTGSATN TKKDATAADV KLTTKLADTD KLEAVADTVD VDAKVKAAET AKKANSADVY	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA LGADVEFGNG DDVKKAATV EADDFKGLGL AALADTDAAL KHAEAFNDIA AAGKAEAAAG TREESDSKFV	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD WNGLARYSYA AIAAAYNNGQ KKVVTNLTKT DATTNALNKL DSLDETNTKA TANTAADKAE RIDGLNATTE	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA GSKQYGNHSG EINGFKAGET VNENKQNVDA GENITTFAEE DEAVKTANEA AVAAKVTDIK KLDTRLASAE	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR RVGVGYRFLE IYDIDEDGTI KVKAAESEIE TKTNIVKIDE KQTAEETKQN ADIATNKDNI KSIADHDTRL
35 40 45	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951 1001 1051 1101 1151 1201 1251 1301 1351	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL NMPHTRLVAG GGGGTGSATN TKKDATAADV KLTTKLADTD KLEAVADTVD VDAKVKAAET AKKANSADVY NGLDKTVSDL	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA LGADVEFGNG DDVKKAATV EADDFKGLGL AALADTDAAL KHAEAFNDIA AAGKAEAAAG TREESDSKFV RKETRQGLAE	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD WNGLARYSYA AIAAAYNNGQ KKVVTNLTKT DATTNALNKL DSLDETNTKA TANTAADKAE RIDGLNATTE QAALSGLFQP	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA GSKQYGNHSG EINGFKAGET VNENKQNVDA GENITTFAEE DEAVKTANEA AVAAKVTDIK KLDTRLASAE YNVGRFNVTA	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR RVGVGYRFLE IYDIDEDGTI KVKAAESEIE TKTNIVKIDE KQTAEETKQN ADIATNKDNI KSIADHDTRL AVGGYKSESA
35 40 45	51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801 851 901 951 1001 1051 1101 1151 1201 1251 1301	VAVTDRDAKI GIVDTGESVG ASFDDEAVIE ATLHIMNTND LFQIANSEEQ IFSTGNDAQA GTEPLEYGSN TAALLLQKYP PASFPFGDFT GKTIIEGGSL DQSGANETVH GAGYLNSTGR GDTLSYYVRR SSATPETVET ATVYADSTAA VEGKMRGSTQ AGIRHDAGDI GGVNVPFAAT LAGLKLSQPL NMPHTRLVAG GGGGTGSATN TKKDATAADV KLTTKLADTD KLEAVADTVD VDAKVKAAET AKKANSADVY NGLDKTVSDL	NAPPPNLHTG SISFPELYGR TEAKPTDIRH ETKNEMMVAA YRQALLDYSG QPNTYALLPF HCGITAMWCL WMSNDNLRTT ADTKGTSDIA VLYGNNKSDM IKGSLQLDGK RVPFLSAAKI GNAARTASAA AAADRTDMPG HADMQGRRLK TVGIAAKTGE GYLKGLFSYG GDLTVEGGLR SDKAVLFATA LGADVEFGNG DDVKKAATV EADDFKGLGL AALADTDAAL KHAEAFNDIA AAGKAEAAAG TREESDSKFV	DFPNPNDAYK KEHGYNENYK VKEIGHIDLV IRNAWVKLGE GDKTDEGIRL YEKDAQKGII SAPYEASVRF LLTTAQDIGA YSFRNDISGT RVETKGALIY GTLYTRLGKL GQDYSFFTNI AHSAPAGLKH IRPYGATFRA AVSDGLDHNG NTTAAATLGM RYKNSISRST YDLLKQDAFA GVERDLNGRD WNGLARYSYA AIAAAYNNGQ KKVVTNLTKT DATTNALNKL DSLDETNTKA TANTAADKAE RIDGLNATTE QAALSGLFQP	NLINLKPAIE NYTAYMRKEA SHIIGGRSVD RGVRIVNNSF MQQSDYGNLS TVAGVDRSGE TRTNPIQIAG VGVDSKFGWG GGLIKKGGSQ NGAASGGSLN LKVDGTAIIG ETDGGLLASL AVEQGGSNLE AAAVQHANAA TGLRVIAQTQ GRSTWSENSA GADEHAEGSV EKGSALGWSG YTVTGGFTGA GSKQYGNHSG EINGFKAGET VNENKQNVDA GENITTFAEE DEAVKTANEA AVAAKVTDIK KLDTRLASAE YNVGRFNVTA	AGYTGRGVEV PEDGGGKDIE GRPAGGIAPD GTTSRAGTAD YHIRNKNMLF KFKREMYGEP TSFSAPIVTG LLDAGKAMNG LQLHGNNTYT SDGIVYLADT GKLYMSARGK DSVEKTAGSE NLMVELDASE DGVRIFNSLA QDGGTWEQGG NAKTDSISLF NGTLMQLGAL NSLTEGTLVG TAATGKTGAR RVGVGYRFLE IYDIDEDGTI KVKAAESEIE TKTNIVKIDE KQTAEETKQN ADIATNKDNI KSIADHDTRL AVGGYKSESA

# <u>ΔG983-961c</u> ATGACTTCTG CGCCCGACTT CAATGCAGGC GGTACCGGTA TCGGCAGCAA CAGCAGAGCA ACAACAGCGA AATCAGCAGC AGTATCTTAC GCCGGTATCA AGAACGAAAT GTGCAAAGAC AGAAGCATGC TCTGTGCCGG TCGGGATGAC GTTGCGGTTA CAGACAGGGA TGCCAAAATC AATGCCCCCC CCCCGAATCT GCATACCGGA GACTTTCCAA ACCCAAATGA CGCATACAAG AATTTGATCA

	251		TGCAATTGAA			
	301		ACACAGGCGA			
	351		AAAGAACACG			
5	401		GAAGGAAGCG ACGATGAGGC			
5	451 501		GTAAAAGAAA			
	551		TTCCGTGGAC			
	601		ACATAATGAA			
	651		ATCCGCAATG			
	701		TAACAGTTTT			
10	751		TAGCCAATTC			
	801		GGTGATAAAA			
	851		CAACCTGTCC			
	901	ATCTTTTCGA	CAGGCAATGA	CGCACAAGCT	CAGCCCAACA	CATATGCCCT
	951	ATTGCCATTT	TATGAAAAAG	ACGCTCAAAA	AGGCATTATC	ACAGTCGCAG
	1001	GCGTAGACCG	CAGTGGAGAA	AAGTTCAAAC	GGGAAATGTA	TGGAGAACCG
15	1051	GGTACAGAAC	CGCTTGAGTA	TGGCTCCAAC	CATTGCGGAA	TTACTGCCAT
	1101		TCGGCACCCT			
	1151		AATTGCCGGA			
	1201		TGCTGCTGCA			
	1251		TTGCTGACGA			
	1301		CGGCTGGGGA			
20	1351	•	TTCCGTTCGG TACTCCTTCC			
	1401 1451		CGGCAGCCAA			
	1501		TTATCGAAGG			
	1551		CGCGTCGAAA			
	1601		CAGCCTGAAC			
	1651		GCGCAAACGA			
25	1701		GGTACGCTGT			
	1751	ACGGTACGGC	GATTATCGGC	GGCAAGCTGT	ACATGTCGGC	ACGCGGCAAG
	1801	GGGGCAGGCT	ATCTCAACAG	TACCGGACGA	CGTGTTCCCT	TCCTGAGTGC
	1851	CGCCAAAATC	GGGCAGGATT	ATTCTTTCTT	CACAAACATC	GAAACCGACG
	1901		GGCTTCCCTC			
30	1951		TGTCCTATTA			
30	2001		GCACATTCCG			
	2051		CAATCTGGAA			
	2101		CACCCGAGAC ATCCGCCCCT			
	2151 2201		GAATGCCGCC			
	2251		ATGCCGACAG			
35	2301		GCCGTATCGG			
	2351		GCAAACCCAA			
	2401		AAATGCGCGG			
	2451	AACCGGCGAA	AATACGACAG	CAGCCGCCAC	ACTGGGCATG	GGACGCAGCA
	2501	CATGGAGCGA	AAACAGTGCA	AATGCAAAAA	CCGACAGCAT	TAGTCTGTTT
	2551		GGCACGATGC			
40	2601		CGCTACAAAA			
	2651		AGGCAGCGTC			
	2701		ACGTTCCGTT			
	2751		TACGACCTGC			
	2801 2851		CTGGAGCGGC TGAAGCTGTC			
45	2901		GGCGTGGAAC			
45	2951		TACCGGCGCG			
	3001		ACACCCGTCT			
	3051		TGGAACGGCT			
	3101		CCACAGCGGA			
	3151		GCACTGGATC			
50	3201		GCCATTGCTG			
	3251		TGGAGAGACC			
	3301	ACCAAAAAAG	ACGCAACTGC	AGCCGATGTT	GAAGCCGACG	ACTTTAAAGG
	3351		AAAAAAGTCG			
	3401		CGTCGATGCC			
	3451		CCAAGTTAGC			
55	3501		GATGCAACCA			
	3551	TAACGACATT	TGCTGAAGAG	ACTAAGACAA	ATATCGTAAA	AATTGATGAA

36	01 AAA	TTAGAAG	CCGTGGCTGA	TACCGTCGAC	AAGCATGCCG	AAGCATTCAA
36	51 CGA	TATCGCC	GATTCATTGG	ATGAAACCAA	CACTAAGGCA	GACGAAGCCG
37	01 TCA	AAACCGC	CAATGAAGCC	AAACAGACGG	CCGAAGAAAC	CAAACAAAAC
5 37	'51 GTC	GATGCCA	AAGTAAAAGC	TGCAGAAACT	GCAGCAGGCA	AAGCCGAAGC
38	01 TGC	CGCTGGC	ACAGCTAATA	CTGCAGCCGA	CAAGGCCGAA	GCTGTCGCTG
38	51 CAA	AAGTTAC	CGACATCAAA	GCTGATATCG	CTACGAACAA	AGATAATATT
39	01 GCT	'AAAAAAG	CAAACAGTGC	CGACGTGTAC	ACCAGAGAAG	AGTCTGACAG
39	51 CAA	ATTTGTC	AGAATTGATG	GTCTGAACGC	TACTACCGAA	AAATTGGACA
40	001 CAC	GCTTGGC	TTCTGCTGAA	AAATCCATTG	CCGATCACGA	TACTCGCCTG
10 40	51 AAC	GGTTTGG	ATAAAACAGT	GTCAGACCTG	CGCAAAGAAA	CCCGCCAAGG
41	.01 CCT	TGCAGAA	CAAGCCGCGC	TCTCCGGTCT	GTTCCAACCT	TACAACGTGG
41	.51 GTC	TCGAGCA	CCACCACCAC	CACCACTGA		
15						
13			GTGIGSNSRA			
1			NAPPPNLHTG SISFPELYGR			
			TEAKPTDIRH			
			ETKNEMMVAA			
			YROALLDYSG			
20		-	OPNTYALLPF			
			HCGITAMWCL			
			WMSNDNLRTT			
		_	ADTKGTSDIA	-		
			VLYGNNKSDM			
-			IKGSLQLDGK			
	_		RVPFLSAAKI			
			GNAARTASAA			
			AAADRTDMPG			
7	751 ATV	YADSTAA	HADMQGRRLK	AVSDGLDHNG	TGLRVIAQTQ	QDGGTWEQGG
8	01 VE	KMRGSTQ	TVGIAAKTGE	NTTAAATLGM	GRSTWSENSA	NAKTDSISLF
30 8	851 AGI	RHDAGDI	GYLKGLFSYG	RYKNSISRST	GADEHAEGSV	NGTLMQLGAL
g	01 GGV	NVPFAAT	GDLTVEGGLR	YDLLKQDAFA	EKGSALGWSG	NSLTEGTLVG
g			SDKAVLFATA			
10	001 NMF	PHTRLVAG	LGADVEFGNG	WNGLARYSYA	GSKQYGNHSG	RVGVGYRFLE
10	)51 GGG	GTGSATN	DDDVKKAATV	AIAAAYNNGQ	EINGFKAGET	IYDIDEDGTI
11			EADDFKGLGL			
			AALADTDAAL			
			KHAEAFNDIA			
			AAGKAEAAAG			
			TREESDSKFV			
13	851 NGI	DKTVSDL	RKETRQGLAE	QAALSGLFQP	YNVGLEHHHH	HH*

∆G741 and hybrids

40

45

50

55

**[0225]** Bactericidal titres generated in response to  $\Delta$ G741 (His-fusion) were measured against various strains, including the homologous 2996 strain:

	2996	MC58	NGH38	F6124	BZ133
∆G741	512	131072	>2048	16384	>2048

[0226] As can be seen, the ΔG741-induced anti-bactericidal titre is particularly high against heterologous strain MC58.
 [0227] ΔG741 was also fused directly in-frame upstream of proteins 961, 961c, 983 and ORF46.1:

	ΔG741-961						
	1		CCGACATCGG	TGCGGGGCTT	GCCGATGCAC	TAACCGCACC	
	51					GATCAGTCCG	
	101					GGAAAAAACT	
5	151					ACGACAAGGT	
	201					CAGCTCATTA	
	201	0.100001110	, dheirimice	, 00011111001		, 01.001 01.1111	
	0.51	000000000000000000000000000000000000000	mcca ca cmmc	CAACMAMACA	77C777CC7	mmcccccmma	
10	251		TGGAGAGTTC AGACCGAGCA				
	301						
	351		CGCCAGTTCA				
	401		GCTTCCCGAA				
	451		ACGATGCCGG				
15	501		GGAAACGGCA				
13	551		GGCCGCCGCC				
	601		GTTCCGTCCT				
	651		TTTGGCGGAA				
	701		AAACGGCATA				
	751		GAGGCACTGG				
20	801		GTGGCCATTG				
	851		AGCTGGAGAG				
	901		AAGACGCAAC				
	951		CTGAAAAAAG				
	1001		AAACGTCGAT				
	1051		CAACCAAGTT CTGGATGCAA				
25	1101		ATTTGCTGAA				
	1151 1201		AAGCCGTGGC				
	1251						
			GCCGATTCAT CGCCAATGAA				
	1301 1351		CCAAAGTAAA				
	1401		GGCACAGCTA				
30	1451		TACCGACATC				
	1501		AAGCAAACAG				
	1551		GTCAGAATTG				
	1601		GGCTTCTGCT				
	1651		TGGATAAAAC				
35	1701		GAACAAGCCG				
	1751		CAATGTAACG				
	1801		TCGGTACCGG				
	1851		GCAGTCGGCA				
	1901	TCGGCGTCAA	TTACGAGTGG	CTCGAGCACC	ACCACCACCA	CCACTGA	
40							
	1		ADALTAPLDH				
	51		KLKNDKVSRF				
	101		SEHSGKMVAK				
45	151		TYTIDFAAKQ				
43	201		AEKGSYSLGI			_	
	251		NDDDVKKAAT				
	301		VEADDFKGLG		_		
	351		DAALADTDAA				
	401		DKHAEAFNDI			_	
50	451		TAAGKAEAAA				
	501		YTREESDSKF				
	551		LRKETRQGLA				
	601	AVAIGTGERF	TENFAAKAGV	AVGTSSGSSA	AYHVGVNYEW	генинини.	

	40744 0					
	<u>ΔG741−9</u>		GGGT GT TTGGG	mcccccccmm	CCCCAMCCAC	maaaaaaaaa
	1		CCGACATCGG			
5	51 101		AAAGACAAAG CGAGAAACTG			
	151		GTGACAGCCT			
	201		GACTTTATCC			
	251		TGGAGAGTTC			
	301		AGACCGAGCA			
	351		CGCCAGTTCA			
10	401		GCTTCCCGAA			
	451		ACGATGCCGG			
	501		GGAAACGGCA			
	551		GGCCGCCGCC			
	601		GTTCCGTCCT			
15	651		TTTGGCGGAA			
10	701		AAACGGCATA			
	751		GAGGCACTGG			
	,01	000010000	0/10/0/10/100		12.00.100.100	
20						
	801		GTGGCCATTG			
	851		AGCTGGAGAG			
	901		AAGACGCAAC			
	951		CTGAAAAAAG			
25	1001		AAACGTCGAT			
	1051		CAACCAAGTT			
	1101		CTGGATGCAA			
	1151		ATTTGCTGAA			
	1201		AAGCCGTGGC			
	1251		GCCGATTCAT			
30	1301		CGCCAATGAA			
	1351		CCAAAGTAAA			
	1401		GGCACAGCTA			
	1451		TACCGACATC			
	1501		AAGCAAACAG			
	1551		GTCAGAATTG			
35	1601		GGCTTCTGCT			
	1651		TGGATAAAAC			
	1701		GAACAAGCCG			CCTTACAACG
	1751	TGGGTCTCGA	GCACCACCAC	CACCACCACT	GA	
40						
	1	MVAADIGAGL	ADALTAPLDH	KDKGLOSLTL	DOSVRKNEKL	KLAAQGAEKT
						QVYKQSHSAL
	101		SEHSGKMVAK			
	151		TYTIDFAAKQ			
45	201		AEKGSYSLGI			
70	251	-	NDDDVKKAAT	_		
	301		VEADDFKGLG		-	
	351		DAALADTDAA			
	401	EKLEAVADTV	DKHAEAFNDI	ADSLDETNTK	ADEAVKTANE	AKQTAEETKQ
	451		TAAGKAEAAA			
50	501	IAKKANSADV	YTREESDSKF	VRIDGLNATT	EKLDTRLASA	EKSIADHDTR
	551		LRKETRQGLA			

	ΔG741-9	<u>83</u>				
	1	ATGGTCGCCG	CCGACATCGG	TGCGGGGCTT	GCCGATGCAC	TAACCGCACC
	51	GCTCGACCAT	AAAGACAAAG	GTTTGCAGTC	TTTGACGCTG	GATCAGTCCG
5	101	TCAGGAAAAA	CGAGAAACTG	AAGCTGGCGG	CACAAGGTGC	GGAAAAAACT
· ·	151	TATGGAAACG	GTGACAGCCT	CAATACGGGC	AAATTGAAGA	ACGACAAGGT
	201	CAGCCGTTTC	GACTTTATCC	GCCAAATCGA	AGTGGACGGG	CAGCTCATTA
	251	CCTTGGAGAG	TGGAGAGTTC	CAAGTATACA	AACAAAGCCA	TTCCGCCTTA
	301	ACCGCCTTTC	AGACCGAGCA	AATACAAGAT	TCGGAGCATT	CCGGGAAGAT
	351	GGTTGCGAAA	CGCCAGTTCA	GAATCGGCGA	CATAGCGGGC	GAACATACAT
10	401	CTTTTGACAA	GCTTCCCGAA	GGCGGCAGGG	CGACATATCG	CGGGACGGCG
	451	TTCGGTTCAG	ACGATGCCGG	CGGAAAACTG	ACCTACACCA	TAGATTTCGC
	501	CGCCAAGCAG	GGAAACGGCA	AAATCGAACA	TTTGAAATCG	CCAGAACTCA
	551	ATGTCGACCT	GGCCGCCGCC	GATATCAAGC	CGGATGGAAA	ACGCCATGCC
	601	GTCATCAGCG	GTTCCGTCCT	TTACAACCAA	GCCGAGAAAG	GCAGTTACTC
	651	CCTCGGTATC	TTTGGCGGAA	AAGCCCAGGA	AGTTGCCGGC	AGCGCGGAAG
15	701	TGAAAACCGT	AAACGGCATA	CGCCATATCG	GCCTTGCCGC	CAAGCAACTC
	751	GAGGGATCCG	GCGGAGGCGG	CACTTCTGCG	CCCGACTTCA	ATGCAGGCGG
	801	TACCGGTATC	GGCAGCAACA	GCAGAGCAAC	AACAGCGAAA	TCAGCAGCAG
	851	TATCTTACGC	CGGTATCAAG	AACGAAATGT	GCAAAGACAG	AAGCATGCTC
	901	TGTGCCGGTC	GGGATGACGT	TGCGGTTACA	GACAGGGATG	CCAAAATCAA
20	951	TGCCCCCCC	CCGAATCTGC	ATACCGGAGA	CTTTCCAAAC	CCAAATGACG
20	1001	CATACAAGAA	TTTGATCAAC	CTCAAACCTG	CAATTGAAGC	AGGCTATACA
	1051	GGACGCGGGG	TAGAGGTAGG	TATCGTCGAC	ACAGGCGAAT	CCGTCGGCAG
	1101	CATATCCTTT	CCCGAACTGT		AGAACACGGC	TATAACGAAA
	1151	ATTACAAAAA	CTATACGGCG	TATATGCGGA	AGGAAGCGCC	TGAAGACGGA
	1201	GGCGGTAAAG	ACATTGAAGC	TTCTTTCGAC	GATGAGGCCG	TTATAGAGAC
25	1251	TGAAGCAAAG	CCGACGGATA	TCCGCCACGT	AAAAGAAATC	GGACACATCG
	1301	ATTTGGTCTC	CCATATTATT		CCGTGGACGG	
	1351	GGCGGTATTG	CGCCCGATGC	GACGCTACAC	ATAATGAATA	CGAATGATGA
	1401	AACCAAGAAC	GAAATGATGG	TTGCAGCCAT	CCGCAATGCA	TGGGTCAAGC
	1451	TGGGCGAACG	TGGCGTGCGC	ATCGTCAATA	ACAGTTTTGG	AACAACATCG
	1501	AGGGCAGGCA	CTGCCGACCT	TTTCCAAATA	GCCAATTCGG	AGGAGCAGTA
30						

	1551			ATTCCGGCGG		
	1601			GATTACGGCA		
	1651			CTTTTCGACA		
5	1701			TGCCATTTTA		
9	1751			GTAGACCGCA		
	1801			TACAGAACCG		
	1851			GGTGCCTGTC		
	1901			CCGATTCAAA		
	1951	GCACCCATCG	TAACCGGCAC	GGCGGCTCTG	CTGCTGCAGA	AATACCCGTG
10	2001	GATGAGCAAC	GACAACCTGC	GTACCACGTT	GCTGACGACG	GCTCAGGACA
	2051	TCGGTGCAGT	CGGCGTGGAC	AGCAAGTTCG	GCTGGGGACT	GCTGGATGCG
	2101	GGTAAGGCCA	TGAACGGACC	CGCGTCCTTT	CCGTTCGGCG	ACTTTACCGC
	2151	CGATACGAAA	GGTACATCCG	ATATTGCCTA	CTCCTTCCGT	AACGACATTT
	2201	CAGGCACGGG	CGGCCTGATC	AAAAAAGGCG	GCAGCCAACT	GCAACTGCAC
	2251	GGCAACAACA	CCTATACGGG	CAAAACCATT	ATCGAAGGCG	GTTCGCTGGT
15	2301	GTTGTACGGC	AACAACAAAT	CGGATATGCG	CGTCGAAACC	AAAGGTGCGC
	2351	TGATTTATAA	CGGGGCGGCA	TCCGGCGGCA	GCCTGAACAG	CGACGGCATT
	2401	GTCTATCTGG	CAGATACCGA	CCAATCCGGC	GCAAACGAAA	CCGTACACAT
	2451	CAAAGGCAGT	CTGCAGCTGG	ACGGCAAAGG	TACGCTGTAC	ACACGTTTGG
	2501	GCAAACTGCT	GAAAGTGGAC	GGTACGGCGA	TTATCGGCGG	CAAGCTGTAC
	2551	ATGTCGGCAC	GCGGCAAGGG	GGCAGGCTAT	CTCAACAGTA	CCGGACGACG
20	2601	TGTTCCCTTC	CTGAGTGCCG	CCAAAATCGG	GCAGGATTAT	TCTTTCTTCA
	2651	CAAACATCGA	AACCGACGGC	GGCCTGCTGG	CTTCCCTCGA	CAGCGTCGAA
	2701	AAAACAGCGG	GCAGTGAAGG	CGACACGCTG	TCCTATTATG	TCCGTCGCGG
	2751	CAATGCGGCA	CGGACTGCTT	CGGCAGCGGC	ACATTCCGCG	CCCGCCGGTC
	2801	TGAAACACGC	CGTAGAACAG	GGCGGCAGCA	ATCTGGAAAA	CCTGATGGTC
	2851	GAACTGGATG	CCTCCGAATC	ATCCGCAACA	CCCGAGACGG	TTGAAACTGC
25	2901	GGCAGCCGAC	CGCACAGATA	TGCCGGGCAT	CCGCCCCTAC	GGCGCAACTT
	2951	TCCGCGCAGC	GGCAGCCGTA	CAGCATGCGA	ATGCCGCCGA	CGGTGTACGC
	3001	ATCTTCAACA	GTCTCGCCGC	TACCGTCTAT	GCCGACAGTA	CCGCCGCCCA
	3051	TGCCGATATG	CAGGGACGCC	GCCTGAAAGC	CGTATCGGAC	GGGTTGGACC
	3101	ACAACGGCAC	GGGTCTGCGC	GTCATCGCGC	AAACCCAACA	GGACGGTGGA
00	3151	ACGTGGGAAC	AGGGCGGTGT	TGAAGGCAAA	ATGCGCGGCA	GTACCCAAAC
30	3201	CGTCGGCATT	GCCGCGAAAA	CCGGCGAAAA	TACGACAGCA	GCCGCCACAC
	3251	TGGGCATGGG	ACGCAGCACA	TGGAGCGAAA	ACAGTGCAAA	TGCAAAAACC
	3301	GACAGCATTA	GTCTGTTTGC	AGGCATACGG	CACGATGCGG	GCGATATCGG
	3351	CTATCTCAAA	GGCCTGTTCT	CCTACGGACG	CTACAAAAAC	AGCATCAGCC
	3401	GCAGCACCGG	TGCGGACGAA	CATGCGGAAG	GCAGCGTCAA	CGGCACGCTG
35	3451	ATGCAGCTGG	GCGCACTGGG	CGGTGTCAAC	GTTCCGTTTG	CCGCAACGGG
00	3501	AGATTTGACG	GTCGAAGGCG	GTCTGCGCTA	CGACCTGCTC	AAACAGGATG
	3551	CATTCGCCGA	AAAAGGCAGT	GCTTTGGGCT	GGAGCGGCAA	CAGCCTCACT
	3601			CGCGGGTCTG		
	3651			CAACGGCGGG		
	3701			GGCGGCTTTA		
40	3751			TATGCCGCAC		
	3801			GCAACGGCTG		
	3851			TACGGCAACC		
	3901			CCACCACCAC		

	1	MVAADTGAGI.	ADALTAPLDH	KDKGLOSLTL	DOSVRKNEKI	KLAAOGAEKT
	51		KLKNDKVSRF			_
	101		SEHSGKMVAK			
5	151		TYTIDFAAKQ			
	201		AEKGSYSLGI			
	251		PDFNAGGTGI			
	301		DRDAKINAPP			
	351	GRGVEVGIVD	TGESVGSISF	PELYGRKEHG	YNENYKNYTA	YMRKEAPEDG
	401	GGKDIEASFD	DEAVIETEAK	PTDIRHVKEI	GHIDLVSHII	GGRSVDGRPA
10	451	GGIAPDATLH	IMNTNDETKN	EMMVAAIRNA	WVKLGERGVR	IVNNSFGTTS
	501	RAGTADLFQI	ANSEEQYRQA	LLDYSGGDKT	DEGIRLMQQS	DYGNLSYHIR
	551	NKNMLFIFST	GNDAQAQPNT	YALLPFYEKD	AOKGIITVAG	VDRSGEKFKR
	601		LEYGSNHCGI			
	651		LLQKYPWMSN			
	701		PFGDFTADTK			
15	751		IEGGSLVLYG			
	801					
			ANETVHIKGS			
	851	MSARGKGAGY	LNSTGRRVPF	LSAAKIGQDY	SEFTNIETDG	GLLASLDSVE
20						
					D7 G7 77117 1750	
	901		SYYVRRGNAA			
	951		PETVETAAAD			
	1001		ADSTAAHADM	-		
	1051	TWEQGGVEGK	MRGSTQTVGI	AAKTGENTTA	AATLGMGRST	WSENSANAKT
25	1101	DSISLFAGIR	HDAGDIGYLK	GLFSYGRYKN	SISRSTGADE	HAEGSVNGTL
23	1151	MQLGALGGVN	VPFAATGDLT	VEGGLRYDLL	KQDAFAEKGS	ALGWSGNSLT
	1201	EGTLVGLAGL	KLSQPLSDKA	VLFATAGVER	DLNGRDYTVT	GGFTGATAAT
	1251		TRLVAGLGAD			
	1301	GYRFLEHHHH		121011011102	THAT DITHOUT Y	1011110011101
	1301	OTKI DDIIIIII	****			
30						
30						
35						
33						
40						
40						
45						
45						
50						
55						

	ΔG741-C	RF46.1				
	1	ATGGTCGCCG	CCGACATCGG	TGCGGGGCTT	GCCGATGCAC	TAACCGCACC
	51	GCTCGACCAT	AAAGACAAAG	GTTTGCAGTC	TTTGACGCTG	GATCAGTCCG
_	101	TCAGGAAAAA	CGAGAAACTG	AAGCTGGCGG	CACAAGGTGC	GGAAAAAACT
5	151	TATGGAAACG	GTGACAGCCT	CAATACGGGC	AAATTGAAGA	ACGACAAGGT
	201	CAGCCGTTTC	GACTTTATCC	GCCAAATCGA	AGTGGACGGG	CAGCTCATTA
	251	CCTTGGAGAG	TGGAGAGTTC	CAAGTATACA	AACAAAGCCA	TTCCGCCTTA
	301	ACCGCCTTTC	AGACCGAGCA	AATACAAGAT	TCGGAGCATT	CCGGGAAGAT
	351	GGTTGCGAAA	CGCCAGTTCA	GAATCGGCGA	CATAGCGGGC	GAACATACAT
10	401	CTTTTGACAA	GCTTCCCGAA	GGCGGCAGGG	CGACATATCG	CGGGACGGCG
10	451	TTCGGTTCAG	ACGATGCCGG	CGGAAAACTG	ACCTACACCA	TAGATTTCGC
	501	CGCCAAGCAG	GGAAACGGCA	AAATCGAACA	TTTGAAATCG	CCAGAACTCA
	551	ATGTCGACCT	GGCCGCCGCC	GATATCAAGC	CGGATGGAAA	ACGCCATGCC
	601			TTACAACCAA		
	651	CCTCGGTATC	TTTGGCGGAA	AAGCCCAGGA	AGTTGCCGGC	AGCGCGGAAG
15	701			CGCCATATCG		
	751			ATCCTCAGAT		
	801			AGCATTTCGA		
	851			CTTGCCGAGC		
	901			GTTGGGCAAC		
	951			ACATTGTCCG		
20	1001			AACCATGCCT		
	1051	GCCGGTAGTC	CCGTTGACGG	ATTTAGCCTT	TACCGCATCC	ATTGGGACGG
	1101	ATACGAACAC	CATCCCGCCG	ACGGCTATGA	CGGGCCACAG	GGCGGCGGCT
	1151	ATCCCGCTCC	CAAAGGCGCG	AGGGATATAT	ACAGCTACGA	CATAAAAGGC
	1201	GTTGCCCAAA	ATATCCGCCT	CAACCTGACC	GACAACCGCA	GCACCGGACA
	1251	ACGGCTTGCC	GACCGTTTCC	ACAATGCCGG	TAGTATGCTG	ACGCAAGGAG
25	1301	TAGGCGACGG	ATTCAAACGC	GCCACCCGAT	ACAGCCCCGA	GCTGGACAGA
	1351	TCGGGCAATG	CCGCCGAAGC	CTTCAACGGC	ACTGCAGATA	TCGTTAAAAA
	1401	CATCATCGGC	GCGGCAGGAG	AAATTGTCGG	CGCAGGCGAT	GCCGTGCAGG
	1451	GCATAAGCGA	AGGCTCAAAC	ATTGCTGTCA	TGCACGGCTT	GGGTCTGCTT
	1501	TCCACCGAAA	ACAAGATGGC	GCGCATCAAC	GATTTGGCAG	ATATGGCGCA
30	1551	ACTCAAAGAC	TATGCCGCAG	CAGCCATCCG	CGATTGGGCA	GTCCAAAACC
30	1601	CCAATGCCGC	ACAAGGCATA	GAAGCCGTCA	GCAATATCTT	TATGGCAGCC
	1651	ATCCCCATCA	AAGGGATTGG	AGCTGTTCGG	GGAAAATACG	GCTTGGGCGG
	1701			AGCGGTCGCA		
	1751	CGAAAGGGAA	ATCCGCCGTC	AGCGACAATT	TTGCCGATGC	GGCATACGCC
	1801			TTCCCGAAAT		
35	1851	GCGTTACGGC	AAAGAAAACA	TCACCTCCTC	AACCGTGCCG	CCGTCAAACG
	1901	GCAAAAATGT	CAAACTGGCA	GACCAACGCC	ACCCGAAGAC	AGGCGTACCG
	1951	TTTGACGGTA	AAGGGTTTCC	GAATTTTGAG	AAGCACGTGA	AATATGATAC
	2001	GCTCGAGCAC	CACCACCACC	ACCACTGA		
40						
40	1	MVAADTGAGI.	ADALTAPLDH	KDKGLQSLTL	DOSVRKNEKI	KLAAOGAEKT
	51			DFIRQIEVDG		
	101			ROFRIGDIAG		
	151			GNGKIEHLKS		
	201			FGGKAQEVAG		
45	251			LDRQHFEPDG		
	301			NIGYIVRFSD		
	351			HPADGYDGPQ		
	401			DRFHNAGSML		
	451			AAGEIVGAGD		
	501			YAAAAIRDWA		
50	551			HPIKRSQMGA		
	601			KENITSSTVP		
	651		KHVKYDTLEH			

# Example 16 - C-terminal fusions ('hybrids') with 287/∆G287

55

**[0228]** According to the invention, hybrids of two proteins A & B may be either  $NH_2$ -A-B-COOH or  $NH_2$ -B-A-COOH. The effect of this difference was investigated using protein 287 either C-terminal (in '287-His' form) or N-terminal (in

ΔG287 form - sequences shown above) to 919, 953 and ORF46.1. A panel of strains was used, including homologous strain 2996. FCA was used as adjuvant:

	287 & 919		287 & 953		287 & ORF46.1	
Strain	∆G287-919	919-287	∆G287-953	953-287	∆G287-46.1	46.1-287
2996	128000	16000	65536	8192	16384	8192
BZ232	256	128	128	<4	<4	<4
1000	2048	<4	<4	<4	<4	<4
MC58	8192	1024	16384	1024	512	128
NGH38	32000	2048	>2048	4096	16384	4096
<b>394</b> /98	4096	32	256	128	128	16
MenA (F6124)	32000	2048	>2048	32	8192	1024
MenC (BZ133)	64000	>8192	>8192	<16	8192	2048

Better bactericidal titres are generally seen with 287 at the N-terminus (in the  $\Delta G$  form)

**[0229]** When fused to protein 961 [NH $_2$ - $\Delta$ G287-961-COOH - sequence shown above], the resulting protein is insoluble and must be denatured and renatured for purification. Following renaturation, around 50% of the protein was found to remain insoluble. The soluble and insoluble proteins were compared, and much better bactericidal titres were obtained with the soluble protein (FCA as adjuvant):

	2996	BZ232	MC58	NGH38	F6124	BZ133
Soluble	65536	128	4096	>2048	>2048	4096
Insoluble	8192	<4	<4	16	n.d.	n.d.

[0230] Titres with the insoluble form were, however, improved by using alum adjuvant instead:

Insoluble	32768	1128	4096	>2048	>2048	2048
-----------	-------	------	------	-------	-------	------

### Example 17 — N-terminal fusions ('hybrids') to 287

5

10

15

20

30

35

40

45

50

**[0231]** Expression of protein 287 as full-length with a C-terminal His-tag, or without its leader peptide but with a C-terminal His-tag, gives fairly low expression levels. Better expression is achieved using a N-terminal GST-fusion.

**[0232]** As an alternative to using GST as an N-terminal fusion partner, 287 was placed at the C-terminus of protein 919 ('919-287'), of protein 953 ('953-287'), and of proteins ORF46.1 ('ORF46.1-287'). In both cases, the leader peptides were deleted, and the hybrids were direct in-frame fusions.

**[0233]** To generate the 953-287 hybrid, the leader peptides of the two proteins were omitted by designing the forward primer downstream from the leader of each sequence; the stop codon sequence was omitted in the 953 reverse primer but included in the 287 reverse primer. For the 953 gene, the 5' and the 3' primers used for amplification included a *Ndel* and a *Bam*HI restriction sites respectively, whereas for the amplification of the 287 gene the 5' and the 3' primers included a *Bam*HI and a *Xho*I restriction sites respectively. In this way a sequential directional cloning of the two genes in pET21b+, using *Ndel-Bam*HI (to clone the first gene) and subsequently *Bam*HI-*Xho*I (to clone the second gene) could be achieved.

[0234] The 919-287 hybrid was obtained by cloning the sequence coding for the mature portion of 287 into the *Xho*l site at the 3'-end of the 919-His clone in pET21b+. The primers used for amplification of the 287 gene were designed for introducing a *Sal*l restriction site at the 5'-and a *Xho*l site at the 3'- of the PCR fragment. Since the cohesive ends produced by the *Sal*l and *Xho*l restriction enzymes are compatible, the 287 PCR product digested with *Sal*l-Xhol could be inserted in the pET21 b-919 clone cleaved with *Xho*l.

[0235] The ORF46.1-287 hybrid was obtained similarly.

[0236] The bactericidal efficacy (homologous strain) of antibodies raised against the hybrid proteins was compared

with antibodies raised against simple mixtures of the component antigens:

	Mixture with 287	Hybrid with 287
919	32000	16000
953	8192	8192
ORF46.1	128	8192

[0237] Data for bactericidal activity against heterologous MenB strains and against serotypes A and C were also obtained for 919-287 and 953-287:

	919		953		ORF46.1	
Strain	Mixture	Hybrid	Mixture	Hybrid	Mixture	Hybrid
MC58	512	1024	512	1024	-	1024
NGH38	1024	2048	2048	4096	-	4096
BZ232	512	128	1024	16	-	-
MenA (F6124)	512	2048	2048	32	-	1024
MenC (C11)	>2048	n.d.	>2048	n.d.	-	n.d.
MenC (BZ133)	>4096	>8192	>4096	<16	-	2048

[0238] Hybrids of ORF46.1 and 919 were also constructed. Best results (four-fold higher titre) were achieved with 919 at the N-terminus.

[0239] Hybrids 919-519His, ORF97-225His and 225-ORF97His were also tested. These gave moderate ELISA fitres and bactericidal antibody responses.

#### Example 18 - the leader peptide from ORF4

[0240] As shown above, the leader peptide of ORF4 can be fused to the mature sequence of other proteins (e.g. proteins 287 and 919). It is able to direct lipidation in *E.coli*.

#### Example 19 - domains in 564

5

15

20

25

30

35

40

45

50

55

**[0241]** The protein '564' is very large (2073aa), and it is difficult to clone and express it in complete form. To facilitate expression, the protein has been divided into four domains, as shown in figure 8 (according to the MC58 sequence):

Domain	Α	В	С	D
Amino Acids	79-360	361-731	732-2044	2045-2073

[0242] These domains show the following homologies:

Domain A shows homology to other bacterial toxins:

gbIAAG03431.1IAE004443\_9probable hemagglutinin [Pseudomonas aeruginosa] (38%) gbIAAC31981.1I (139897) HecA [Pectobacterium chrysanthemi] (45%) embICAA36409.1I (X52156) filamentous hemagglutinin [Bordetella pertussis] (31%) gbIAAC79757.1I (AF057695) large supernatant protein1 [Haemophilus ducreyi] (26%) gbIAAA25657.1I (M30186) HpmA precursor [Proteus mirabilis] (29%)

- Domain B shows no homology, and is specific to 564.
- Domain C shows homology to:

gbIAAF84995.1IAE004032 HA-like secreted protein [Xylella fastidiosa] (33%) gbIAAG05850.1IAE004673 hypothetical protein [Pseudomonas aeruginosa] (27%) gbIAAF68414.1AF237928 putative FHA [Pasteurella multocisìda] (23%) gbIAAC79757.1I(AF057695) large supernatant protein1 [Haemophilus ducreyi] (23%) pirllS21010 FHA B precursor [Bordetella pertussis] (20%)

Domain D shows homology to other bacterial toxins:

gbIAAF84995.1IAE004032 14 HA-like secreted protein [Xylella fastidiosa] (29%)

[0243] Using the MC58 strain sequence, good intracellular expression of 564ab was obtained in the form of GSTfusions (no purification) and his-tagged protein; this domain-pair was also expressed as a lipoprotein, which showed moderate expression in the outer membrane/supernatant fraction.

[0244] The b domain showed moderate intracellular expression when expressed as a his-tagged product (no purification), and good expression as a GST-fusion.

[0245] The c domain showed good intracellular expression as a GST-fusion, but was insoluble. The d domain showed moderate intracellular expression as a his-tagged product (no purification). The cd protein domain-pair showed moderate intracellular expression (no purification) as a GST-fusion.

[0246] Good bactericidal assay titres were observed using the c domain and the bc pair.

#### Example 20 - the 919 leader peptide

[0247] The 20mer leader peptide from 919 is discussed in example 1 above: MKKYLFRAAL YGIAAAILAA

[0248] As shown in example 1, deletion of this leader improves heterologous expression, as does substitution with the ORF4 leader peptide. The influence of the 919 leader on expression was investigated by fusing the coding sequence to the PhoC reporter gene from Morganella morganii [Thaller et al. (1994) Microbiology 140:1341-1350]. The construct was cloned in the pET21-b plasmid between the Ndel and Xhol sites (Figure 9):

> 1 MKKYLFRAAL YGIAAAILAA AIPAGNDATT KPDLYYLKNE QAIDSLKLLP PPPEVGSIQF LNDQAMYEKG RMLRNTERGK QAQADADLAA GGVATAFSGA 101 FGYPITEKDS PELYKLLTNM IEDAGDLATR SAKEHYMRIR PFAFYGTETC 151 NTKDQKKLST NGSYPSGHTS IGWATALVLA EVNPANQDAI LERGYQLGQS

> RVICGYHWQS DVDAARIVGS AAVATLHSDP AFQAQLAKAK QEFAQKSQK\*

[0249] The level of expression of PhoC from this plasmid is >200-fold lower than that found for the same construct but containing the native PhoC signal peptide. The same result was obtained even after substitution of the T7 promoter with the E.coli Plac promoter. This means that the influence of the 919 leader sequence on expression does not depend on the promoter used.

[0250] In order to investigate if the results observed were due to some peculiarity of the 919 signal peptide nucleotide sequence (secondary structure formation, sensitivity to RNAases, etc.) or to protein instability induced by the presence of this signal peptide, a number of mutants were generated. The approach used was a substitution of nucleotides of the 919 signal peptide sequence by cloning synthetic linkers containing degenerate codons. In this way, mutants were obtained with nucleotide and/or amino acid substitutions.

[0251] Two different linkers were used, designed to produce mutations in two different regions of the 919 signal peptide sequence, in the first 19 base pairs (L1) and between bases 20-36 (S1).

L1: 5' T ATG AAa/g TAc/t c/tTN TTt/c a/cGC GCC GCC CTG TAC GGC ATC GCC GCC GCC ATC CTC GCC GCC GCG ATC CC 3'

S1: 5' T ATG AAA AAA TAC CTA TTC CGa/g GCN GCN c/tTa/g TAc/t GGc/g ATC GCC GCC GCC ATC CTC GCC GCC GCG ATC CC 3'

[0252] The alignment of some of the mutants obtained is given below.

52

10

5

30

35

40

50

#### L1 mutants:

	9L1-a	ATGAAGAAGTACCTTTTCAGCGCCGCC~~~~~~~~~~~~~
	9L1-e	ATGAAAAAATACTTTTTCCGCGCCGCC~~~~~~~~~~~~~
	9L1-d	ATGAAAAAATACTTTTTCCGCGCCGCC~~~~~~~~~~~~~
5	9L1-f	ATGAAAAAATATCTCTTTAGCGCCGCCCTGTACGGCATCGCCGCCGCCATCCTCGCCGCC
	919sp	ATGAAAAAATACCTATTCCGCGCCGCCCTGTACGGCATCGCCGCCATCCTCGCCGCC
	9L1a	MKKYLFSAA~~~~~~
	9L1e	MKKYFFRAA~~~~~~~
10	9L1d	MKKYFFRAA~~~~~~~
.0	9L1f	MKKYLFSAALYGIAAAILAA
	91 <b>9</b> sp	MKKYLFRAALYGIAAAILAA ( $i.e.$ native signal peptide)

# 15 S1 mutants:

9S1-e	ATGAAAAATACCTATTCATCGCCGCCGCCATCCTCGCCGCC
9S1-c	ATGAAAAAATACCTATTCCGAGCTGCCCAATACGGCATCGCCGCCGCCATCCTCGCCGCC
9S1-b	ATGAAAAAATACCTATTCCGGGCCGCCCAATACGGCATCGCCGCCGCCATCCTCGCCGCC
9 <b>S</b> 1-i	ATGAAAAATACCTATTCCGGGCGGCTTTGTACGGGATCGCCGCCGCCATCCTCGCCGCC
919sp	ATGAAAAAATACCTATTCCGCGCCCCCTGTACGGCATCGCCGCCCCCCCC
9S1e	MKKYLFIAAAILAA
9S1c	MKKYLFRAAQYGIAAAILAA
9S1b	MKKYLFRAAQYGIAAAILAA
9S1i	MKKYLFRAALYGIAAAILAA
919sp	MKKYLFRAALYGIAAAILAA

[0253] As shown in the sequences alignments, most of the mutants analysed contain in-frame deletions which were unexpectedly produced by the host cells.

[0254] Selection of the mutants was performed by transforming *E. coli* BL21(DE3) cells with DNA prepared from a mixture of L1 and S1 mutated clones. Single transformants were screened for high PhoC activity by streaking them onto LB plates containing 100 μg/ml ampicillin, 50μg/ml methyl green, 1 mg/ml PDP (phenolphthaleindiphosphate). On this medium PhoC-producing cells become green (Figure 10).

**[0255]** A quantitative analysis of PhoC produced by these mutants was carried out in liquid medium using pNPP as a substrate for PhoC activity. The specific activities measured in cell extracts and supernatants of mutants grown in liquid medium for 0, 30, 90, 180 min. were:

0,00 4,44 172,05 83,25 3,11

> 36,63 28,86 14,43

26,64 142,08

#### **CELL EXTRACTS**

#### 40 [0256]

20

25

30

		0	30	90
	control	0,00	0,00	0,00
45	9phoC	1,11	1,11	3,33
	9S1e	102,12	111,00	149,85
	9L1a	206,46	111,00	94,35
50	9L1d	5,11	4,77	4,00
	9L1f	27,75	94,35	82,14
	9S1b	156,51	111,00	72,15
	9S1c	72,15	33,30	21,09
55	9S1i	156,51	83,25	55,50
	phoCwt	194,25	180,93	149,85

#### **SUPERNATANTS**

#### [0257]

5		
10		
15		

20

30

35

0 30 90 1801 control 0,00 0,00 0,00 0,00 9phoC 0,33 0,00 0,00 0,00 9S1e 0.89 0.11 0,22 0.44 9L1a 4,88 5.99 5,99 7,22 9L1d 0,11 0,11 0,11 0,11 9L1f 0,22 0,11 0,11 0,11 9S1b 1,44 1.44 1.44 1,67 9S1c 0,78 0,56 0,67 0,44 9S1i 0,44 0,22 0,22 0,78 43,29 87,69 phoCwt 34,41 177,60

**[0258]** Some of the mutants produce high amounts of PhoC and in particular, mutant 9L1a can secrete PhoC in the culture medium. This is noteworthy since the signal peptide sequence of this mutant is only 9 amino acids long. This is the shortest signal peptide described to date.

#### Example 21— C-terminal deletions of Maf-related proteins

[0259] MafB-related proteins include 730, ORF46 and ORF29.

[0260] The 730 protein from MC58 has the following sequence:

1	VKPLRRLTNL	LAACAVAAAA	LIQPALAADL	AQDPFITDNA	QRQHYEPGGK
51	YHLFGDPRGS	VSDRTGKINV	IQDYTHQMGN	LLIQQANING	TIGYHTRFSG
101	HGHEEHAPFD	NHAADSASEE	KGNVDEGFTV	YRLNWEGHEH	HPADAYDGPK
151	GGNYPKPTGA	RDEYTYHVNG	TARSIKLNPT	DTRSIRQRIS	DNYSNLGSNF
201	SDRADEANRK	MFEHNAKLDR	WGNSMEFING	VAAGALNPFI	SAGEALGIGD
251	ILYGTRYAID	KAAMRNIAPL	PAEGKFAVIG	GLGSVAGFEK	NTREAVDRWI
301	QENPNAAETV	EAVFNVAAAA	KVAKLAKAAK	PGKAAVSGDF	ADSYKKKLAL

40

- 351 SDSARQLYQN AKYREALDIH YEDLIRRKTD GSSKFINGRE IDAVTNDALI
- 401 QAKRTISAID KPKNFLNQKN RKQIKATIEA ANQQGKRAEF WFKYGVHSQV
- 451 KSYIESKGGI VKTGLGD\*

45

55

[0261] The leader peptide is underlined.

[0262] 730 shows similar features to ORF46 (see example 8 above):

th

- as for Orf46, the conservation of the 730 sequence among MenB, MenA and gonococcus is high (>80%) only for the N-terminal portion. The C-terminus, from ~340, is highly divergent.
- its predicted secondary structure contains a hydrophobic segment spanning the central region of the molecule (aa. 227-247).
- expression of the full-length gene in *E. coli* gives very low yields of protein. Expression from tagged or untagged constructs where the signal peptide sequence has been omitted has a toxic effect on the host cells. In other words, the presence of the full-length mature protein in the cytoplasm is highly toxic for the host cell while its translocation to the periplasm (mediated by the signal peptide) has no detectable effect on cell viability. This "intracellular toxicity" of 730 is particularly high since clones for expression of the leaderless 730 can only be obtained at very low frequency using a *recA* genetic background (*E. coli* strains: HB101 for cloning; HMS174(DE3) for expression).

**[0263]** To overcome this toxicity, a similar approach was used for 730 as described in example 8 for ORF46. Four C-terminal truncated forms were obtained, each of which is well expressed. All were obtained from intracellular expression of His-tagged leaderless 730.

[0264] Form A consists of the N-terminal hydrophilic region of the mature protein (aa. 28-226). This was purified as a soluble His-tagged product, having a higher-than-expected MW.

[0265] Form B extends to the end of the region conserved between serogroups (aa. 28-340). This was purified as an insoluble His-tagged product.

[0266] The C-terminal truncated forms named C1 and C2 were obtained after screening for clones expressing high levels of 730-His clones in strain HMS174(DE3). Briefly, the pET21b plasmid containing the His-tagged sequence coding for the full-length mature 730 protein was used to transform the *recA* strain HMS 174(DE3). Transformants were obtained at low frequency which showed two phenotypes: large colonies and very small colonies. Several large and small colonies were analysed for expression of the 730-His clone. Only cells from large colonies over-expressed a protein recognised by anti-730A antibodies. However the protein over-expressed in different clones showed differences in molecular mass. Sequencing of two of the clones revealed that in both cases integration of an *E. coli* IS sequence had occurred within the sequence coding for the C terminal region of 730. The two integration events have produced in-frame fusion with 1 additional codon in the case of C1, and 12 additional codons in the case of C2 (Figure 11). The resulting "mutant" forms of 730 have the following sequences:

# 730-C1 (due to an IS1 insertion - figure 11A) 1 MADLAQDPFI TDNAQRQHYE PGGKYHLFGD PRGSVSDRTG KINVIQDYTH 51 QMGNLLIQQA NINGTIGYHT RFSGHGHEEH APFDNHAADS ASEEKGNVDE 101 GFTVYRLNWE GHEHHPADAY DGPKGGNYPK PTGARDEYTY HVNGTARSIK 151 LNPTDTRSIR QRISDNYSNL GSNFSDRADE ANRKMFEHNA KLDRWGNSME 201 FINGVAAGAL NPFISAGEAL GIGDILYGTR YAIDKAAMRN IAPLPAEGKF

251 AVIGGLGSVA GFEKNTREAV DRWIQENPNA AETVEAVFNV AAAAKVAKLA 301 KAAKPGKAAV SGDFADSYKK KLALSDSARQ LYQNAKYREA LDIHYEDLIR

351 RKTDGSSKFI NGREIDAVTN DALIQAR\*

[0267] The additional amino acid produced by the insertion is underlined.

#### 730-C2 (due to an IS5 insertion - Figure 11B)

1 MADLAQDPFI TDNAQRQHYE PGGKYHLFGD PRGSVSDRTG KINVIQDYTH
51 QMGNLLIQQA NINGTIGYHT RFSGHGHEEH APFDNHAADS ASEEKGNVDE
101 GFTVYRLNWE GHEHHPADAY DGPKGGNYPK PTGARDEYTY HVNGTARSIK
151 LNPTDTRSIR QRISDNYSNL GSNFSDRADE ANRKMFEHNA KLDRWGNSME
201 FINGVAAGAL NPFISAGEAL GIGDILYGTR YAIDKAAMRN IAPLPAEGKF
251 AVIGGLGSVA GFEKNTREAV DRWIQENPNA AETVEAVFNV AAAAKVAKLA
301 KAAKPGKAAV SGDFADSYKK KLALSDSARQ LYQNAKYREA LGKVRISGEI
351 LLG\*

[0268] The additional amino acids produced by the insertion are underlined.

**[0269]** In conclusion, intracellular expression of the 730-C1 form gives very high level of protein and has no toxic effect on the host cells, whereas the presence of the native C-terminus is toxic. These data suggest that the "intracellular toxicity" of 730 is associated with the C-terminal 65 amino acids of the protein.

**[0270]** Equivalent truncation of ORF29 to the first 231 or 368 amino acids has been performed, using expression with or without the leader peptide (amino acids 1-26; deletion gives cytoplasmic expression) and with or without a His-tag.

#### Example 22 - domains in 961

20

25

30

35

40

50

55

**[0271]** As described in example 9 above, the GST-fusion of 961 was the best-expressed in *E.coli*. To improve expression, the protein was divided into domains (figure 12).

**[0272]** The domains of 961 were designed on the basis of YadA (an adhesin produced by *Yersinia* which has been demonstrated to be an adhesin localized on the bacterial surface that forms oligomers that generate surface projection [Hoiczyk et al. (2000) EMBO J 19:5989-99]) and are: leader peptide, head domain, coiled-coil region (stalk), and membrane anchor domain.

[0273] These domains were expressed with or without the leader peptide, and optionally fused either to C-terminal His-tag or to N-terminal GST. *E.coli* clones expressing different domains of 961 were analyzed by SDS-PAGE and

western blot for the production and localization of the expressed protein, from over-night (o/n) culture or after 3 hours induction with IPTG. The results were:

	Total lysate (Western Blot)	Periplasm (Western Blot)	Supernatant (Western Blot)	OMV SDS-PAGE
961 (o/n) 961 (IPTG)	- +/-	-	-	
961-L (o/n) 961-L (IPTG)	+ +	-	-	+ +
961c-L (o/n) 961 c-L (IPTG)	- +	- +	-+	
961Δ <sub>1</sub> -L (o/n) 961Δ <sub>1</sub> -L (IPTG)	-+	-	-	+

[0274] The results show that in E. coli:

5

10

15

20

25

30

35

40

45

55

- 961-L is highly expressed and localized on the outer membrane. By western blot analysis two specific bands have been detected: one at ~45kDa (the predicted molecular weight) and one at ~180kDa, indicating that 961-L can form oligomers. Additionally, these aggregates are more expressed in the over-night culture (without IPTG induction). OMV preparations of this clone were used to immunize mice and serum was obtained. Using overnight culture (predominantly by oligomeric form) the serum was bactericidal; the IPTG-induced culture (predominantly monomeric) was not bactericidal.
- 961 $\Delta_1$ -L (with a partial deletion in the anchor region) is highly expressed and localized on the outer membrane, but does not form oligomers;
- the 961c-L (without the anchor region) is produced in soluble form and exported in the supernatant.

[0275] Titres in ELISA and in the serum bactericidal assay using His-fusions were as follows:

	ELISA	Bactericidal
961a (aa 24-268)	24397	4096
961b (aa 269-405)	7763	64
961c-L	29770	8192
961c (2996)	30774	>65536
961c (MC58)	33437	16384
961d	26069	>65536

**[0276]** E.coli clones expressing different forms of 961 (961, 961-L, 961 $\Delta_1$ -L and 961c-L) were used to investigate if the 961 is an adhesin (c.f. YadA). An adhesion assay was performed using (a) the human epithelial cells and (b) E.coli clones after either over-night culture or three hours IPTG induction. 961-L grown over-night (961 $\Delta_1$ -L) and IPTG-induced 961c-L (the clones expressing protein on surface) adhere to human epithelial cells.

[0277] 961c was also used in hybrid proteins (see above). As 961 and its domain variants direct efficient expression, they are ideally suited as the N-terminal portion of a hybrid protein.

#### Example 23 — further hybrids

[0278] Further hybrid proteins of the invention are shown below (see also Figure 14). These are advantageous when compared to the individual proteins:

	ORF46.1	-741				
	1		TGGCAAACGA	TTCTTTTATC	CGGCAGGTTC	TCGACCGTCA
	51					AGGGGGGAAC
	101			ATCGGATTGG		
5	151			ACAGGCGGCC		
	201			ACGGGCACGA		
	251					CGTTGACGGA
	301	TTTAGCCTTT	ACCGCATCCA	TTGGGACGGA	TACGAACACC	ATCCCGCCGA
	351					AAAGGCGCGA
10	401			ATAAAAGGCG		
10	451			CACCGGACAA		
	501			CGCAAGGAGT		
	551			CTGGACAGAT		
	601	TTCAACGGCA	CTGCAGATAT	CGTTAAAAAC	ATCATCGGCG	CGGCAGGAGA
	651			CCGTGCAGGG		
15	701			GGTCTGCTTT		
	751			TATGGCGCAA		
	801			TCCAAAACCC		
	851			ATGGCAGCCA		
	901			CTTGGGCGGC		
	951			TCGCATTGCC		
20	1001			GCATACGCCA		
	1051			CTTGGAGCAG		
	1101			CGTCAAACGG		
	1151			GGCGTACCGT		
	1201			ATATGATACG		
	1251			GGCTTGCCGA		
25	1301			CAGTCTTTGA		
	1351			GGCGGCACAA		
	1401			CGGGCAAATT		
	1451			ATCGAAGTGG		
	1501			ATACAAACAA		
	1551			AAGATTCGGA		
30	1601			GGCGACATAG		
	1001	CGAAACGCCA	GIICAGAAIC	GGCGACATAG	COOCGAACA	INCATCITII
35						
33	1651			CAGGGCGACA		
	1701			AACTGACCTA		
	1751			GAACATTTGA		
	1801			CAAGCCGGAT		
	1851			ACCAAGCCGA		
40	1901			CAGGAAGTTG		
	1951			TATCGGCCTT	GCCGCCAAGC	AACTCGAGCA
	2001	CCACCACCAC	CACCACTGA			
45	1	MSDLANDSFI	ROVLDROHFE	PDGKYHLFGS	RGELAERSGH	IGLGKIOSHO
,,,	51			FSDHGHEVHS		
	101			GPQGGGYPAP		
	151			SMLTQGVGDG		_
	201	_		AGDAVQGISE		
	251			DWAVQNPNAA		
50	301			MGAIALPKGK		
	351			TVPPSNGKNV		
	401	_		IGAGLADALT	<del></del>	
	451			SLNTGKLKND		
	501			EQIQDSEHSG		
	551			AGGKLTYTID		
55	601			VLYNQAEKGS		
	651		AAKQLEHHHH		-DESTI SOM	×
	031	1 1110111111111111111111111111111111111		****		

	ORF46.1	-961				
	1		TGGCAAACGA	TTCTTTTATC	CGGCAGGTTC	TCGACCGTCA
	51		CCCGACGGGA			
5	101		CAGCGGCCAT			
	151		TGATGATTCA			
	201		TTTTCCGATC			
	251		ACATTCCGAT			
	301		ACCGCATCCA	•••		
	351	CGGCTATGAC	GGGCCACAGG	GCGGCGGCTA	TCCCGCTCCC	AAAGGCGCGA
10	401	GGGATATATA	CAGCTACGAC	ATAAAAGGCG	TTGCCCAAAA	TATCCGCCTC
	451	AACCTGACCG	ACAACCGCAG	CACCGGACAA	CGGCTTGCCG	ACCGTTTCCA
	501	CAATGCCGGT	AGTATGCTGA	CGCAAGGAGT	AGGCGACGGA	TTCAAACGCG
	551	CCACCCGATA	CAGCCCCGAG	CTGGACAGAT	CGGGCAATGC	CGCCGAAGCC
	601	TTCAACGGCA	CTGCAGATAT	CGTTAAAAAC	ATCATCGGCG	CGGCAGGAGA
	651	AATTGTCGGC	GCAGGCGATG	CCGTGCAGGG	CATAAGCGAA	GGCTCAAACA
15	701	TTGCTGTCAT	GCACGGCTTG	GGTCTGCTTT	CCACCGAAAA	CAAGATGGCG
	751	CGCATCAACG	ATTTGGCAGA	TATGGCGCAA	CTCAAAGACT	ATGCCGCAGC
	801	AGCCATCCGC	GATTGGGCAG	TCCAAAACCC	CAATGCCGCA	CAAGGCATAG
	851	AAGCCGTCAG	CAATATCTTT	ATGGCAGCCA	TCCCCATCAA	AGGGATTGGA
	901	GCTGTTCGGG	GAAAATACGG	CTTGGGCGGC	ATCACGGCAC	ATCCTATCAA
20	951	GCGGTCGCAG	ATGGGCGCGA	TCGCATTGCC	GAAAGGGAAA	TCCGCCGTCA
20	1001	GCGACAATTT	TGCCGATGCG	GCATACGCCA	AATACCCGTC	CCCTTACCAT
	1051	TCCCGAAATA	TCCGTTCAAA	CTTGGAGCAG	CGTTACGGCA	AAGAAAACAT
	1101	CACCTCCTCA	ACCGTGCCGC	CGTCAAACGG	CAAAAATGTC	AAACTGGCAG
	1151	ACCAACGCCA	CCCGAAGACA	GGCGTACCGT	TTGACGGTAA	AGGGTTTCCG
	1201		AGCACGTGAA			
25	1251	CACAAACGAC	GACGATGTTA	AAAAAGCTGC	CACTGTGGCC	ATTGCTGCTG
	1301	CCTACAACAA	TGGCCAAGAA	ATCAACGGTT	TCAAAGCTGG	AGAGACCATC
	1351	TACGACATTG	ATGAAGACGG	CACAATTACC	AAAAAAGACG	CAACTGCAGC
	1401	CGATGTTGAA	GCCGACGACT	TTAAAGGTCT	GGGTCTGAAA	AAAGTCGTGA
	1451	CTAACCTGAC	CAAAACCGTC	AATGAAAACA	AACAAAACGT	CGATGCCAAA
	1501	GTAAAAGCTG	CAGAATCTGA	AATAGAAAAG	TTAACAACCA	AGTTAGCAGA
30	1551		GCTTTAGCAG			
	1601	ACGCCTTGAA	TAAATTGGGA	GAAAATATAA	CGACATTTGC	TGAAGAGACT
	1651	AAGACAAATA	TCGTAAAAAT	TGATGAAAAA	TTAGAAGCCG	TGGCTGATAC
	1701	CGTCGACAAG	CATGCCGAAG	CATTCAACGA	TATCGCCGAT	TCATTGGATG
	1751	AAACCAACAC	TAAGGCAGAC	GAAGCCGTCA	AAACCGCCAA	TGAAGCCAAA
0.5	1801	CAGACGGCCG	AAGAAACCAA	ACAAAACGTC	GATGCCAAAG	TAAAAGCTGC
35	1851	AGAAACTGCA	GCAGGCAAAG	CCGAAGCTGC	CGCTGGCACA	GCTAATACTG
	1901	CAGCCGACAA	GGCCGAAGCT	GTCGCTGCAA	AAGTTACCGA	CATCAAAGCT
	1951	GATATCGCTA	CGAACAAAGA	TAATATTGCT	AAAAAAGCAA	ACAGTGCCGA
	2001	CGTGTACACC	AGAGAAGAGT	CTGACAGCAA	ATTTGTCAGA	ATTGATGGTC
40						
,,,						
	2051		TACCGAAAAA			
	2101		ATCACGATAC			
	2151		AAAGAAACCC			
45	2201	CCGGTCTGTT	CCAACCTTAC	AACGTGGGTC	GGTTCAATGT	AACGGCTGCA
	2251		ACAAATCCGA			
	2301	CTTTACCGAA	AACTTTGCCG	CCAAAGCAGG	CGTGGCAGTC	GGCACTTCGT
	2351		CGCAGCCTAC		TCAATTACGA	GTGGCTCGAG
	2401	CACCACCACC	ACCACCACTG	A		

	1	MSDLANDSFI	RQVLDRQHFE	PDGKYHLFGS	RGELAERSGH	IGLGKIQSHQ
	51	LGNLMIQQAA	IKGNIGYIVR	FSDHGHEVHS	PFDNHASHSD	SDEAGSPVDG
	101	FSLYRIHWDG	YEHHPADGYD	GPQGGGYPAP	KGARDIYSYD	IKGVAQNIRL
5	151	NLTDNRSTGQ	RLADRFHNAG	SMLTQGVGDG	FKRATRYSPE	LDRSGNAAEA
3	201	FNGTADIVKN	IIGAAGEIVG	AGDAVQGISE	GSNIAVMHGL	GLLSTENKMA
	251	RINDLADMAQ	LKDYAAAAIR	DWAVQNPNAA	QGIEAVSNIF	MAAIPIKGIG
	301	AVRGKYGLGG	ITAHPIKRSQ	MGAIALPKGK	SAVSDNFADA	AYAKYPSPYH
	351	SRNIRSNLEQ	RYGKENITSS	TVPPSNGKNV	KLADQRHPKT	GVPFDGKGFP
	401	NFEKHVKYDT	GSGGGGATND	DDVKKAATVA	IAAAYNNGQE	INGFKAGETI
10	451	YDIDEDGTIT	KKDATAADVE	ADDFKGLGLK	KVVTNLTKTV	NENKQNVDAK
	501	VKAAESEIEK	LTTKLADTDA	ALADTDAALD	ATTNALNKLG	ENITTFAEET
	551	KTNIVKIDEK	LEAVADTVDK	HAEAFNDIAD	SLDETNTKAD	EAVKTANEAK
	601	QTAEETKQNV	DAKVKAAETA	AGKAEAAAGT	ANTAADKAEA	VAAKVTDIKA
	651	DIATNKDNIA	KKANSADVYT	REESDSKFVR	IDGLNATTEK	LDTRLASAEK
	701	SIADHDTRLN	GLDKTVSDLR	KETRQGLAEQ	AALSGLFQPY	NVGRFNVTAA
15	751	VGGYKSESAV	AIGTGFRFTE	NFAAKAGVAV	GTSSGSSAAY	HVGVNYEWLE
	801	нннннн*				

# 20 ORF46.1-961c

20	OR	F40.1-3	<u> 1016</u>				
		1 7	ATGTCAGATT	TGGCAAACGA	TTCTTTTATC	CGGCAGGTTC	TCGACCGTCA
		51 (	GCATTTCGAA	CCCGACGGGA	AATACCACCT	ATTCGGCAGC	AGGGGGGAAC
		101 7	TTGCCGAGCG	CAGCGGCCAT	ATCGGATTGG	GAAAAATACA	AAGCCATCAG
	•	151 7	TTGGGCAACC	TGATGATTCA	ACAGGCGGCC	ATTAAAGGAA	ATATCGGCTA
	:	201 (	CATTGTCCGC	TTTTCCGATC	ACGGGCACGA	AGTCCATTCC	CCCTTCGACA
25	;	251 <i>I</i>	ACCATGCCTC	ACATTCCGAT	TCTGATGAAG	CCGGTAGTCC	CGTTGACGGA
				ACCGCATCCA			
	:			GGGCCACAGG			
	•	401 (	GGGATATATA	CAGCTACGAC	ATAAAAGGCG	TTGCCCAAAA	TATCCGCCTC
	•	451 <i>I</i>	AACCTGACCG	ACAACCGCAG	CACCGGACAA	CGGCTTGCCG	ACCGTTTCCA
20	!			AGTATGCTGA			
30	!			CAGCCCCGAG			
				CTGCAGATAT			
				GCAGGCGATG			
			TTGCTGTCAT	GCACGGCTTG	GGTCTGCTTT	CCACCGAAAA	CAAGATGGCG
				ATTTGGCAGA			
35			AGCCATCCGC	GATTGGGCAG	TCCAAAACCC	CAATGCCGCA	CAAGGCATAG
				CAATATCTTT			
				GAAAATACGG			
				ATGGGCGCGA			
	10			TGCCGATGCG			
				TCCGTTCAAA			
40				ACCGTGCCGC			
				CCCGAAGACA			
				AGCACGTGAA			
				GACGATGTTA			
				TGGCCAAGAA			
				ATGAAGACGG			
45		-		GCCGACGACT			
	_			CAAAACCGTC			
				CAGAATCTGA			
				GCTTTAGCAG			
				TAAATTGGGA			
<i></i> 0				TCGTAAAAAT			
50				CATGCCGAAG			
				TAAGGCAGAC			
				AAGAAACCAA			
	18	851 <i>I</i>	AGAAACTGCA	GCAGGCAAAG	CCGAAGCTGC	CGCTGGCACA	GCTAATACTG

5	1901 1951 2001 2051 2101 2151 2201 2251	GATATCGCTA CGTGTACACC TGAACGCTAC TCCATTGCCG AGACCTGCGC	GGCCGAAGCT CGAACAAGA AGAGAAGAGT TACCGAAAAA ATCACGATAC AAAGAAACCC CCAACCTTAC	TAATATTGCT CTGACAGCAA TTGGACACAC TCGCCTGAAC GCCAAGGCCT	AAAAAGCAA ATTTGTCAGA GCTTGGCTTC GGTTTGGATA TGCAGAACAA	ACAGTGCCGA ATTGATGGTC TGCTGAAAAA AAACAGTGTC GCCGCGCTCT
10						
15	1 51 101 151 201 251 301	LGNLMIQQAA FSLYRIHWDG NLTDNRSTGQ FNGTADIVKN RINDLADMAQ AVRGKYGLGG	RQVLDRQHFE IKGNIGYIVR YEHHPADGYD RLADRFHNAG IIGAAGEIVG LKDYAAAAIR ITAHPIKRSQ	FSDHGHEVHS GPQGGGYPAP SMLTQGVGDG AGDAVQGISE DWAVQNPNAA MGAIALPKGK	PFDNHASHSD KGARDIYSYD FKRATRYSPE GSNIAVMHGL QGIEAVSNIF SAVSDNFADA	SDEAGSPVDG IKGVAQNIRL LDRSGNAAEA GLLSTENKMA MAAIPIKGIG AYAKYPSPYH
20	351 401 451 501 551 601	NFEKHVKYDT YDIDEDGTIT VKAAESEIEK KTNIVKIDEK	RYGKENITSS GSGGGGATND KKDATAADVE LTTKLADTDA LEAVADTVDK DAKVKAAETA	DDVKKAATVA ADDFKGLGLK ALADTDAALD HAEAFNDIAD	IAAAYNNGQE KVVTNLTKTV ATTNALNKLG SLDETNTKAD	INGFKAGETI NENKQNVDAK ENITTFAEET EAVKTANEAK
25	651 701 751	DIATNKDNIA	KKANSADVYT GLDKTVSDLR	REESDSKFVR	IDGLNATTEK	LDTRLASAEK
30						
35						
40						
45						
50						
55						

	961-ORF	46.1				
	1		ACGACGACGA	ТСТТААААА	GCTGCCACTG	TGGCCATTGC
	51				CGGTTTCAAA	
-	101				TTACCAAAAA	
5	151				GGTCTGGGTC	
	201				AAACAAACAA	
	251				AAAAGTTAAC	
	301				GATGCCGCTC	
	351				TATAACGACA	
10	401				AAAAATTAGA	
10	451				AACGATATCG	
	501				CGTCAAAACC	
	551				ACGTCGATGC	
	601				GCTGCCGCTG	
	651				TGCAAAAGTT	
15	701				TTGCTAAAAA	
	701 751				AGCAAATTTG	
					CACACGCTTG	
	801					
	851				TGAACGGTTT	
	901 951				GGCCTTGCAG GGGTCGGTTC	
20	1001				CAGTCGCCAT	
	1001				GCAGGCGTGG	
	1101				CGGCGTCAAT	
	1151				ACGATTCTTT	
	1201				GGGAAATACC	
	1251				CCATATCGGA	
25	1301				TTCAACAGGC	
	1351				GATCACGGGC	
	1401				CGATTCTGAT	
	1451				TCCATTGGGA	
	1501				CAGGGCGGCG	
30	1551				CGACATAAAA	
30	1601				GCAGCACCGG	
	1651				CTGACGCAAG	
	1701				CGAGCTGGAC	
	1751				ATATCGTTAA	
	1801				GATGCCGTGC	
35	1851				CTTGGGTCTG	
	1901				CAGATATGGC	
	1301	7444107410711	CCCCCCCTTC	711007111100	0.10.11.11	
40	1051	CA CHAMCCCC	CACCACCCAM	CCCCCAMMCC	CCACMCCAAA	A CCCCA A MCC
	1951				GCAGTCCAAA CTTTATGGCA	
	2001					
	2051				ACGGCTTGGG	
	2101				GCGATCGCAT	
	2151				TGCGGCATAC	
45	2201				CAAACTTGGA	
	2251				CCGCCGTCAA	
	2301				GACAGGCGTA	
	2351				TGAAATATGA	TACGUTUGAG
	2401	CACCACCACC	ACCACCACTG	A		

5 10	1 51 101 151 201 251 301 351 401 451 501 551 601 651 701 751 801	AADVEADDFK ADTDAALADT DTVDKHAEAF AAETAAGKAE ADVYTREESD VSDLRKETRQ FRFTENFAAK VLDRQHFEPD GNIGYIVRFS HHPADGYDGP ADRFHNAGSM GAAGEIVGAG DYAAAAIRDW AHPIKRSQMG	GLGLKKVVTN DAALDATTNA NDIADSLDET AAAGTANTAA SKFVRIDGLN GLAEQAALSG AGVAVGTSSG GKYHLFGSRG DHGHEVHSPF QGGGYPAPKG LTQGVGDGFK DAVQGISEGS AVQNPNAAQG AIALPKGKSA	NNGQEINGFK LTKTVNENKQ LNKLGENITT NTKADEAVKT DKAEAVAAKV ATTEKLDTRL LFQPYNVGRF SSAAYHVGVN ELAERSGHIG DNHASHSDSD ARDIYSYDIK RATRYSPELD NIAVMHGLGL IEAVSNIFMA VSDNFADAAY ADQRHPKTGV	NVDAKVKAAE FAEETKTNIV ANEAKQTAEE TDIKADIATN ASAEKSIADH NVTAAVGGYK YEWGSGGGGS LGKIQSHQLG EAGSPVDGFS GVAQNIRLNL RSGNAAEAFN LSTENKMARI AIPIKGIGAV AKYPSPYHSR	SEIEKLTTKL KIDEKLEAVA TKQNVDAKVK KDNIAKKANS DTRLNGLDKT SESAVAIGTG DLANDSFIRQ NLMIQQAAIK LYRIHWDGYE TDNRSTGQRL GTADIVKNII NDLADMAQLK RGKYGLGGIT NIRSNLEQRY
	961-741					
20	1			TGTTAAAAAA		
	51			AAGAAATCAA		
	101 151			GACGGCACAA CGACTTTAAA		
				CCGTCAATGA		
	201					
	251			TCTGAAATAG AGCAGATACT		
25	301					
	351			TGGGAGAAAA		
	401			AAAATTGATG		
	451			CGAAGCATTC		
	501			CAGACGAAGC		
	551			ACCAAACAAA		
30	601			CAAAGCCGAA		
	651			AAGCTGTCGC		
	701			AAAGATAATA		
	751			AGAGTCTGAC		
	801			AAAAATTGGA		
0.5	851			GATACTCGCC		
35	901			AACCCGCCAA		
	951			CTTACAACGT		
	1001			TCCGAATCGG		
	1051			TGCCGCCAAA		
	1101			CCTACCATGT		
40	1151			GCCGCCGACA		
40	1201			CCATAAAGAC		
	1251	GCTGGATCAG	TCCGTCAGGA	AAAACGAGAA	ACTGAAGCTG	GCGGCACAAG
	1301	GTGCGGAAAA	AACTTATGGA	AACGGTGACA	GCCTCAATAC	GGGCAAATTG
	1351			TTTCGACTTT		
	1401			AGAGTGGAGA		
45	1451	GCCATTCCGC	CTTAACCGCC	TTTCAGACCG	AGCAAATACA	AGATTCGGAG
,•	1501	CATTCCGGGA	AGATGGTTGC	GAAACGCCAG	TTCAGAATCG	GCGACATAGC
	1551	GGGCGAACAT	ACATCTTTTG	ACAAGCTTCC	CGAAGGCGGC	AGGGCGACAT
	1601	ATCGCGGGAC	GGCGTTCGGT	TCAGACGATG	CCGGCGGAAA	ACTGACCTAC
	1651	ACCATAGATT	TCGCCGCCAA	GCAGGGAAAC	GGCAAAATCG	AACATTTGAA
	1701			ACCTGGCCGC		
50	1751			AGCGGTTCCG		
	1801			TATCTTTGGC		
<i>55</i>		0000		000000		1 mac + + · = =
· <del>-</del>	1851					ATCGGCCTTG
	1901	CCGCCAAGCA	ACTCGAGCAC	CACCACCACC	ACCACTGA	

5	1 51 101 151 201 251 301 351 401	AADVEADDFK ADTDAALADT DTVDKHAEAF AAETAAGKAE ADVYTREESD VSDLRKETRQ FRFTENFAAK	AATVAIAAAY GLGLKKVVTN DAALDATTNA NDIADSLDET AAAGTANTAA SKFVRIDGLN GLAEQAALSG AGVAVGTSSG KGLQSLTLDQ	LTKTVNENKQ LNKLGENITT NTKADEAVKT DKAEAVAAKV ATTEKLDTRL LFQPYNVGRF SSAAYHVGVN	NVDAKVKAAE FAEETKTNIV ANEAKQTAEE TDIKADIATN ASAEKSIADH NVTAAVGGYK YEWGSGGGGV	SEIEKLTTKL KIDEKLEAVA TKQNVDAKVK KDNIAKKANS DTRLNGLDKT SESAVAIGTG AADIGAGLAD
10	451 501 551 601	KNDKVSRFDF HSGKMVAKRQ TIDFAAKQGN	IRQIEVDGQL FRIGDIAGEH GKIEHLKSPE GKAQEVAGSA	ITLESGEFQV TSFDKLPEGG LNVDLAAADI	YKQSHSALTA RATYRGTAFG KPDGKRHAVI	FQTEQIQDSE SDDAGGKLTY SGSVLYNQAE
15						
20						
25						
30						
35						
40						
45						
50						
55						

	061 000					
	961-983 1	7 mcccc7 c7 7	7 CC7 CC7 CC7	m/cmmn n n n n n n	CCTCCCACTC	mcccca mmcc
			ACGACGACGA			
	51		AACAATGGCC			
5	101		CATTGATGAA			
	151		TTGAAGCCGA			
	201		CTGACCAAAA			
	251		AGCTGCAGAA			
	301		ATGCCGCTTT			
	351		TTGAATAAAT			
10	401		AAATATCGTA			
	451		ACAAGCATGC			
	501		AACACTAAGG			
	551		GGCCGAAGAA			
	601		CTGCAGCAGG			
45	651		GACAAGGCCG			
15	701		CGCTACGAAC			
	751		ACACCAGAGA			
	801		GCTACTACCG			
	851		TGCCGATCAC			
	901		TGCGCAAAGA			
20	951		CTGTTCCAAC			
20	1001		CGGCTACAAA			
	1051		CCGAAAACTT			
	1101		TCTTCCGCAG			
	1151	GATCCGGCGG	AGGCGGCACT	TCTGCGCCCG	ACTTCAATGC	AGGCGGTACC
	1201	GGTATCGGCA	GCAACAGCAG	AGCAACAACA	GCGAAATCAG	CAGCAGTATC
25	1251		ATCAAGAACG			
20	1301	CCGGTCGGGA	TGACGTTGCG	GTTACAGACA	GGGATGCCAA	AATCAATGCC
	1351	CCCCCCGA	ATCTGCATAC	CGGAGACTTT	CCAAACCCAA	ATGACGCATA
	1401	CAAGAATTTG	ATCAACCTCA	AACCTGCAAT	TGAAGCAGGC	TATACAGGAC
	1451	GCGGGGTAGA	GGTAGGTATC	GTCGACACAG	GCGAATCCGT	CGGCAGCATA
	1501	TCCTTTCCCG	AACTGTATGG	CAGAAAAGAA	CACGGCTATA	ACGAAAATTA
30	1551	CAAAAACTAT	ACGGCGTATA	TGCGGAAGGA	AGCGCCTGAA	GACGGAGGCG
	1601	GTAAAGACAT	TGAAGCTTCT	TTCGACGATG	AGGCCGTTAT	AGAGACTGAA
	1651	GCAAAGCCGA	CGGATATCCG	CCACGTAAAA	GAAATCGGAC	ACATCGATTT
	1701	GGTCTCCCAT	ATTATTGGCG	GGCGTTCCGT	GGACGGCAGA	CCTGCAGGCG
	1751	GTATTGCGCC	CGATGCGACG	CTACACATAA	TGAATACGAA	TGATGAAACC
	1801	AAGAACGAAA	TGATGGTTGC	AGCCATCCGC	AATGCATGGG	TCAAGCTGGG
35	1851	CGAACGTGGC	GTGCGCATCG	TCAATAACAG	TTTTGGAACA	ACATCGAGGG
	1901	CAGGCACTGC	CGACCTTTTC	CAAATAGCCA	ATTCGGAGGA	GCAGTACCGC
	1951	CAAGCGTTGC	TCGACTATTC	CGGCGGTGAT	AAAACAGACG	AGGGTATCCG
	2001	CCTGATGCAA	CAGAGCGATT	ACGGCAACCT	GTCCTACCAC	ATCCGTAATA
	2051		TTTCATCTTT			
	2101		CCCTATTGCC			
40	2151		GCAGGCGTAG			
	2201		ACCGGGTACA			
	2251		CCATGTGGTG			
	2301		ACAAACCCGA			
	2351		CGGCACGGCG			
	-001	COLLICOTARC	0000100000	2310100100	_ JOI.OI II II II I	2300130/110

	2401	AGCAACGACA	ACCTGCGTAC	CACGTTGCTG	ACGACGGCTC	AGGACATCGG
	2451			AGTTCGGCTG		
	2501	AGGCCATGAA	CGGACCCGCG	TCCTTTCCGT	TCGGCGACTT	TACCGCCGAT
5	2551	ACGAAAGGTA	CATCCGATAT	TGCCTACTCC	TTCCGTAACG	ACATTTCAGG
	2601	CACGGGCGGC	CTGATCAAAA	AAGGCGGCAG	CCAACTGCAA	CTGCACGGCA
	2651	ACAACACCTA	TACGGGCAAA	ACCATTATCG	AAGGCGGTTC	GCTGGTGTTG
	2701	TACGGCAACA	ACAAATCGGA	TATGCGCGTC	GAAACCAAAG	GTGCGCTGAT
	2751	TTATAACGGG	GCGGCATCCG	GCGGCAGCCT	GAACAGCGAC	GGCATTGTCT
	2801	ATCTGGCAGA	TACCGACCAA	TCCGGCGCAA	ACGAAACCGT	ACACATCAAA
10	2851	GGCAGTCTGC	AGCTGGACGG	CAAAGGTACG	CTGTACACAC	GTTTGGGCAA
	2901	ACTGCTGAAA	GTGGACGGTA	CGGCGATTAT	CGGCGGCAAG	CTGTACATGT
	2951	CGGCACGCGG	CAAGGGGGCA	GGCTATCTCA	ACAGTACCGG	ACGACGTGTT
	3001	CCCTTCCTGA	GTGCCGCCAA	AATCGGGCAG	GATTATTCTT	TCTTCACAAA
	3051	CATCGAAACC	GACGGCGGCC	TGCTGGCTTC	CCTCGACAGC	GTCGAAAAAA
15	3101	CAGCGGGCAG	TGAAGGCGAC	ACGCTGTCCT	ATTATGTCCG	TCGCGGCAAT
15	3151	GCGGCACGGA	CTGCTTCGGC	AGCGGCACAT	TCCGCGCCCG	CCGGTCTGAA
	3201	ACACGCCGTA	GAACAGGGCG	GCAGCAATCT	GGAAAACCTG	ATGGTCGAAC
	3251	TGGATGCCTC	CGAATCATCC	GCAACACCCG	AGACGGTTGA	AACTGCGGCA
	3301	GCCGACCGCA	CAGATATGCC	GGGCATCCGC	CCCTACGGCG	CAACTTTCCG
	3351	CGCAGCGGCA	GCCGTACAGC	ATGCGAATGC	CGCCGACGGT	GTACGCATCT
20	3401	TCAACAGTCT	CGCCGCTACC	GTCTATGCCG	ACAGTACCGC	CGCCCATGCC
	3451	GATATGCAGG	GACGCCGCCT	GAAAGCCGTA	TCGGACGGGT	TGGACCACAA
	3501	CGGCACGGGT	CTGCGCGTCA	TCGCGCAAAC	CCAACAGGAC	GGTGGAACGT
	3551	GGGAACAGGG	CGGTGTTGAA	GGCAAAATGC	GCGGCAGTAC	CCAAACCGTC
	3601			CGAAAATACG		
	3651			GCGAAAACAG		
25	3701			ATACGGCACG		
	3751	CTCAAAGGCC	TGTTCTCCTA	CGGACGCTAC	AAAAACAGCA	TCAGCCGCAG
	3801			CGGAAGGCAG		
	3851			GTCAACGTTC		
	3901			GCGCTACGAC		
	3951			TGGGCTGGAG		
30	4001			GGTCTGAAGC		
	4051			GGCGGGCGTG		
	4101			GCTTTACCGG		
	4151			CCGCACACCC		
	4201			CGGCTGGAAC		
0.5	4251			GCAACCACAG		GGCGTAGGCT
35	4301	ACCGGTTCCT	CGAGCACCAC	CACCACCACC	ACTGA	

	1	MATNDDDVKK	AATVAIAAAY	NNGQEINGFK	AGETIYDIDE	DGTITKKDAT
	51	AADVEADDFK	GLGLKKVVTN	LTKTVNENKQ	NVDAKVKAAE	SEIEKLTTKL
	101	ADTDAALADT	DAALDATTNA	LNKLGENITT	FAEETKTNIV	KIDEKLEAVA
	151			NTKADEAVKT		
5	201			DKAEAVAAKV		
_	251			ATTEKLDTRL		
	301			LFOPYNVGRF		
	351	_	-	SSAAYHVGVN		
	401			IKNEMCKDRS		
10	451			INLKPAIEAG		
10	501			TAYMRKEAPE		
	551			IIGGRSVDGR		
	601	KNEMMVAAIR	NAWVKLGERG	VRIVNNSFGT	TSRAGTADLF	QIANSEEQYR
	651	QALLDYSGGD	KTDEGIRLMQ	QSDYGNLSYH	IRNKNMLFIF	STGNDAQAQP
	701	NTYALLPFYE	KDAQKGIITV	AGVDRSGEKF	KREMYGEPGT	EPLEYGSNHC
	751	GITAMWCLŞA	PYEASVRFTR	TNPIQIAGTS	FSAPIVTGTA	ALLLOKYPWM
15	801	SNDNLRTTLL	TTAODIGAVG	VDSKFGWGLL	DAGKAMNGPA	SFPFGDFTAD
	851			LIKKGGSQLQ		
	901			AASGGSLNSD		
	951			VDGTAIIGGK		
	1001	_		DGGLLASLDS		
20	1051			EQGGSNLENL		
	1101			AVQHANAADG		
	1151	DMQGRRLKAV	SDGLDHNGTG	LRVIAQTQQD	GGTWEQGGVE	GKMRGSTQTV
	1201	GIAAKTGENT	TAAATLGMGR	STWSENSANA	KTDSISLFAG	IRHDAGDIGY
	1251	LKGLFSYGRY	KNSISRSTGA	DEHAEGSVNG	TLMQLGALGG	VNVPFAATGD
	1301	LTVEGGLRYD	LLKODAFAEK	GSALGWSGNS	LTEGTLVGLA	GLKLSQPLSD
25			-			
	1351	KAVLFATAGV	ERDLNGRDYT	VTGGFTGATA	ATGKTGARNM	PHTRLVAGLG
	1401	ADVEFGNGWN	GLARYSYAGS	KQYGNHSGRV	GVGYRFLEHH	нннн*
30				~		
35						
33						
40						
45						
50						
55						

#### 961c-ORF46.1 ATGGCCACAA ACGACGACGA TGTTAAAAAA GCTGCCACTG TGGCCATTGC TGCTGCCTAC AACAATGGCC AAGAAATCAA CGGTTTCAAA GCTGGAGAGA 5 CCATCTACGA CATTGATGAA GACGGCACAA TTACCAAAAA AGACGCAACT GCAGCCGATG TTGAAGCCGA CGACTTTAAA GGTCTGGGTC TGAAAAAAGT CGTGACTAAC CTGACCAAAA CCGTCAATGA AAACAAACAA AACGTCGATG CCAAAGTAAA AGCTGCAGAA TCTGAAATAG AAAAGTTAAC AACCAAGTTA 251 GCAGACACTG ATGCCGCTTT AGCAGATACT GATGCCGCTC TGGATGCAAC CACCAACGCC TTGAATAAAT TGGGAGAAAA TATAACGACA TTTGCTGAAG 351 10 AGACTAAGAC AAATATCGTA AAAATTGATG AAAAATTAGA AGCCGTGGCT 401 GATACCGTCG ACAAGCATGC CGAAGCATTC AACGATATCG CCGATTCATT 451 GGATGAAACC AACACTAAGG CAGACGAAGC CGTCAAAACC GCCAATGAAG 501 551 CCAAACAGAC GGCCGAAGAA ACCAAACAAA ACGTCGATGC CAAAGTAAAA GCTGCAGAAA CTGCAGCAGG CAAAGCCGAA GCTGCCGCTG GCACAGCTAA 601 TACTGCAGCC GACAAGGCCG AAGCTGTCGC TGCAAAAGTT ACCGACATCA 651 15 AAGCTGATAT CGCTACGAAC AAAGATAATA TTGCTAAAAA AGCAAACAGT 701 751 GCCGACGTGT ACACCAGAGA AGAGTCTGAC AGCAAATTTG TCAGAATTGA TGGTCTGAAC GCTACTACCG AAAAATTGGA CACACGCTTG GCTTCTGCTG 801 AAAAATCCAT TGCCGATCAC GATACTCGCC TGAACGGTTT GGATAAAACA GTGTCAGACC TGCGCAAAGA AACCCGCCAA GGCCTTGCAG AACAAGCCGC 851 901 GCTCTCCGGT CTGTTCCAAC CTTACAACGT GGGTGGATCC GGAGGAGGAG 951 1001 GATCAGATTT GGCAAACGAT TCTTTTATCC GGCAGGTTCT CGACCGTCAG 1051 CATTTCGAAC CCGACGGGAA ATACCACCTA TTCGGCAGCA GGGGGGAACT 1101 TGCCGAGCGC AGCGGCCATA TCGGATTGGG AAAAATACAA AGCCATCAGT TGGGCAACCT GATGATTCAA CAGGCGGCCA TTAAAGGAAA TATCGGCTAC 1151 ATTGTCCGCT TTTCCGATCA CGGGCACGAA GTCCATTCCC CCTTCGACAA 1201 25 1251 CCATGCCTCA CATTCCGATT CTGATGAAGC CGGTAGTCCC GTTGACGGAT TTAGCCTTTA CCGCATCCAT TGGGACGGAT ACGAACACCA TCCCGCCGAC GGCTATGACG GGCCACAGGG CGGCGGCTAT CCCGCTCCCA AAGGCGCGAG 1351 GGATATATAC AGCTACGACA TAAAAGGCGT TGCCCAAAAT ATCCGCCTCA 1401 ACCTGACCGA CAACCGCAGC ACCGGACAAC GGCTTGCCGA CCGTTTCCAC 1451 1501 AATGCCGGTA GTATGCTGAC GCAAGGAGTA GGCGACGGAT TCAAACGCGC 30 CACCCGATAC AGCCCCGAGC TGGACAGATC GGGCAATGCC GCCGAAGCCT 1551 TCAACGGCAC TGCAGATATC GTTAAAAACA TCATCGGCGC GGCAGGAGAA 1601 1651 ATTGTCGGCG CAGGCGATGC CGTGCAGGGC ATAAGCGAAG GCTCAAACAT 1701 TGCTGTCATG CACGGCTTGG GTCTGCTTTC CACCGAAAAC AAGATGGCGC GCATCAACGA TTTGGCAGAT ATGGCGCAAC TCAAAGACTA TGCCGCAGCA 1751 GCCATCCGCG ATTGGGCAGT CCAAAACCCC AATGCCGCAC AAGGCATAGA 1801 35 1851 AGCCGTCAGC AATATCTTTA TGGCAGCCAT CCCCATCAAA GGGATTGGAG 1901 CTGTTCGGGG AAAATACGGC TTGGGCGGCA TCACGGCACA TCCTATCAAG CGGTCGCAGA TGGGCGCGAT CGCATTGCCG AAAGGGAAAT CCGCCGTCAG 1951 CGACAATTTT GCCGATGCGG CATACGCCAA ATACCCGTCC CCTTACCATT 2001 CCCGAAATAT CCGTTCAAAC TTGGAGCAGC GTTACGGCAA AGAAAACATC 2051 2101 ACCTCCTCAA CCGTGCCGCC GTCAAACGGC AAAAATGTCA AACTGGCAGA 40 CCAACGCCAC CCGAAGACAG GCGTACCGTT TGACGGTAAA GGGTTTCCGA 2151 2201 ATTTTGAGAA GCACGTGAAA TATGATACGC TCGAGCACCA CCACCACCAC 2251 CACTGA

45

50

55

	1	MATNDDDVKK	AATVAIAAAY	NNGQEINGFK	AGETIYDIDE	DGTITKKDAT
	51	AADVEADDFK	GLGLKKVVTN	LTKTVNENKQ	NVDAKVKAAE	SEIEKLTTKL
	101	ADTDAALADT	DAALDATTNA	LNKLGENITT	FAEETKTNIV	KIDEKLEAVA
F	151	DTVDKHAEAF	NDIADSLDET	NTKADEAVKT	ANEAKQTAEE	TKQNVDAKVK
5	201	AAETAAGKAE	AAAGTANTAA	DKAEAVAAKV	TDIKADIATN	KDNIAKKANS
	251	ADVYTREESD	SKFVRIDGLN	ATTEKLDTRL	ASAEKSIADH	DTRLNGLDKT
	301	VSDLRKETRQ	GLAEQAALSG	LFQPYNVGGS	${\tt GGGGSDLAND}$	SFIRQVLDRQ
	351	HFEPDGKYHL	FGSRGELAER	SGHIGLGKIQ	SHQLGNLMIQ	QAAIKGNIGY
	401	IVRFSDHGHE	VHSPFDNHAS	HSDSDEAGSP	VDGFSLYRIH	WDGYEHHPAD
10	451	GYDGPQGGGY	PAPKGARDIY	SYDIKGVAQN	IRLNLTDNRS	TGQRLADRFH
10	501	NAGSMLTQGV	GDGFKRATRY	SPELDRSGNA	AEAFNGTADI	VKNIIGAAGE
	551	IVGAGDAVQG	ISEGSNIAVM	HGLGLLSTEN	KMARINDLAD	MAQLKDYAAA
	601	AIRDWAVQNP	NAAQGIEAVS	NIFMAAIPIK	GIGAVRGKYG	LGGITAHPIK
	651	RSQMGAIALP	KGKSAVSDNF	ADAAYAKYPS	PYHSRNIRSN	LEQRYGKENI
	701	TSSTVPPSNG	KNVKLADQRH	PKTGVPFDGK	GFPNFEKHVK	YDTLEHHHHH
15						

751 H\*

20

#### 961c-741

	1	ATGGCCACAA	ACGACGACGA	TGTTAAAAAA	GCTGCCACTG	TGGCCATTGC
	51	TGCTGCCTAC	AACAATGGCC	AAGAAATCAA	CGGTTTCAAA	GCTGGAGAGA
25	101	CCATCTACGA	CATTGATGAA	GACGGCACAA	TTACCAAAAA	AGACGCAACT
	151	GCAGCCGATG	TTGAAGCCGA	CGACTTTAAA	GGTCTGGGTC	TGAAAAAAGT
	201	CGTGACTAAC	CTGACCAAAA	CCGTCAATGA	AAACAAACAA	AACGTCGATG
	251	CCAAAGTAAA	AGCTGCAGAA	TCTGAAATAG	AAAAGTTAAC	AACCAAGTTA
	301	GCAGACACTG	ATGCCGCTTT	AGCAGATACT	GATGCCGCTC	TGGATGCAAC
	351	CACCAACGCC	TTGAATAAAT	TGGGAGAAAA	TATAACGACA	TTTGCTGAAG
30	401	AGACTAAGAC	AAATATCGTA	AAAATTGATG	AAAAATTAGA	AGCCGTGGCT
	451	GATACCGTCG	ACAAGCATGC	CGAAGCATTC	AACGATATCG	CCGATTCATT
	501	GGATGAAACC	AACACTAAGG	CAGACGAAGC	CGTCAAAACC	GCCAATGAAG
	551	CCAAACAGAC	GGCCGAAGAA	ACCAAACAAA	ACGTCGATGC	CAAAGTAAAA
	601	GCTGCAGAAA	CTGCAGCAGG	CAAAGCCGAA	GCTGCCGCTG	GCACAGCTAA
35	651	TACTGCAGCC	GACAAGGCCG	AAGCTGTCGC	TGCAAAAGTT	ACCGACATCA
55	701	AAGCTGATAT	CGCTACGAAC	AAAGATAATA	TTGCTAAAAA	AGCAAACAGT
	751	GCCGACGTGT	ACACCAGAGA	AGAGTCTGAC	AGCAAATTTG	TCAGAATTGA
	801	TGGTCTGAAC	GCTACTACCG	AAAAATTGGA	CACACGCTTG	GCTTCTGCTG
	851			GATACTCGCC		
	901	GTGTCAGACC	TGCGCAAAGA	AACCCGCCAA	GGCCTTGCAG	AACAAGCCGC
40	951	GCTCTCCGGT	CTGTTCCAAC	CTTACAACGT	GGGTGGATCC	GGAGGGGGTG
	1001	GTGTCGCCGC	CGACATCGGT	GCGGGGCTTG	CCGATGCACT	AACCGCACCG
	1051	CTCGACCATA	AAGACAAAGG	TTTGCAGTCT	TTGACGCTGG	ATCAGTCCGT
	1101			AGCTGGCGGC		
	1151			AATACGGGCA		
	1201			CCAAATCGAA		
45	1251			AAGTATACAA		
	1301			ATACAAGATT	- +	
	1351			AATCGGCGAC		
	1401			GCGGCAGGGC		
	1451			GGAAAACTGA		
	1501			AATCGAACAT		
50	1551			ATATCAAGCC		
	1601			TACAACCAAG		
	1651			AGCCCAGGAA		
	1701			GCCATATCGG	CCTTGCCGCC	AAGCAACTCG
	1751	AGCACCACCA	CCACCACCAC	TGA		

5	1 51 101 151 201 251 301 351 401 451 501 551	AADVEADDFK ADTDAALADT DTVDKHAEAF AAETAAGKAE ADVYTREESD VSDLRKETRQ LDHKDKGLQS SRFDFIRQIE VAKRQFRIGD AKQGNGKIEH	GLGLKKVVTN DAALDATTNA NDIADSLDET AAAGTANTAA SKFVRIDGLN GLAEQAALSG LTLDQSVRKN VDGQLITLES IAGEHTSFDK LKSPELNVDL	NNGQEINGFK LTKTVNENKQ LNKLGENITT NTKADEAVKT DKAEAVAAKV ATTEKLDTRL LFQPYNVGGS EKLKLAAQGA GEFQVYKQSH LPEGGRATYR AAADIKPDGK NGIRHIGLAA	NVDAKVKAAE FAEETKTNIV ANEAKQTAEE TDIKADIATN ASAEKSIADH GGGGVAADIG EKTYGNGDSL SALTAFQTEQ GTAFGSDDAG RHAVISGSVL	SEIEKLTTKL KIDEKLEAVA TKQNVDAKVK KDNIAKKANS DTRLNGLDKT AGLADALTAP NTGKLKNDKV IQDSEHSGKM GKLTYTIDFA YNQAEKGSYS
15						
20	961c-98 1 51 101 151 201 251 301 351	ATGGCCACAA TGCTGCCTAC CCATCTACGA GCAGCCGATG CGTGACTAAC CCAAAGTAAA GCAGACACTG	AACAATGGCC CATTGATGAA TTGAAGCCGA CTGACCAAAA AGCTGCAGAA ATGCCGCTTT	TGTTAAAAA AAGAAATCAA GACGGCACAA CGACTTTAAA CCGTCAATGA TCTGAAATAG AGCAGATACT TGGGAGAAAA	CGGTTTCAAA TTACCAAAAA GGTCTGGGTC AAACAAACAA AAAAGTTAAC GATGCCGCTC	GCTGGAGAGA AGACGCAACT TGAAAAAAGT AACGTCGATG AACCAAGTTA TGGATGCAAC
25	401 451	GATACCGTCG	ACAAGCATGC	AAAATTGATG CGAAGCATTC	AACGATATCG	CCGATTCATT
<i>30</i>	501	GGATGAAACC	AACACTAAGG	CAGACGAAGC	CGTCMAAACC	GCCANIGANG
40						
45						
50						
55						

	551	CCAAACAGAC	GGCCGAAGAA	ACCAAACAAA	ACGTCGATGC	CAAAGTAAAA
	601	GCTGCAGAAA	CTGCAGCAGG	CAAAGCCGAA	GCTGCCGCTG	GCACAGCTAA
	651				TGCAAAAGTT	
5	701				TTGCTAAAAA	
5	751				AGCAAATTTG	
	801				CACACGCTTG	
	851 901				TGAACGGTTT GGCCTTGCAG	
	951				GGGTGGATCC	
	1001				GTACCGGTAT	
10	1051				GTATCTTACG	
	1101				CTGTGCCGGT	
	1151				ATGCCCCCC	
	1201	CATACCGGAG	ACTTTCCAAA	CCCAAATGAC	GCATACAAGA	ATTTGATCAA
	1251				AGGACGCGGG	
	1301				GCATATCCTT	
15	1351				AATTACAAAA	
	1401				AGGCGGTAAA	
	1451				CTGAAGCAAA	
	1501 1551				GATTTGGTCT AGGCGGTATT	
	1601				AAACCAAGAA	
20	1651				CTGGGCGAAC	
20	1701				GAGGGCAGGC	
	1751				ACCGCCAAGC	
	1801				ATCCGCCTGA	
	1851	CGATTACGGC	AACCTGTCCT	ACCACATCCG	TAATAAAAAC	ATGCTTTTCA
	1901				AGCCCAACAC	
25	1951				GGCATTATCA	
	2001				GGAAATGTAT	
	2051				ATTGCGGAAT	
	2101				GTCCGTTTCA CGCACCCATC	
	2151 2201				GGATGAGCAA	
	2251				ATCGGTGCAG	
30	2301				GGGTAAGGCC	
	2351				CCGATACGAA	
	2401	GATATTGCCT	ACTCCTTCCG	TAACGACATT	TCAGGCACGG	GCGGCCTGAT
	2451	CAAAAAAGGC	GGCAGCCAAC	TGCAACTGCA	CGGCAACAAC	ACCTATACGG
	2501				TGTTGTACGG	
35	2551				CTGATTTATA	
35	2601				TGTCTATCTG	
	2651				TCAAAGGCAG	
	2701 2751				GGCAAACTGC CATGTCGGCA	
	2801				GTGTTCCCTT	
	2851				ACAAACATCG	
40	2901				AAAAACAGCG	
	2951	GCGACACGCT	GTCCTATTAT	GTCCGTCGCG	GCAATGCGGC	ACGGACTGCT
	3001	TCGGCAGCGG	CACATTCCGC	GCCCGCCGGT	CTGAAACACG	CCGTAGAACA
	3051				CGAACTGGAT	
	3101				CGGCAGCCGA	
	3151				TTCCGCGCAG	
45	3201				CATCTTCAAC	
	3251				ATGCCGATAT CACAACGGCA	
	3301 3351				AACGTGGGAA	
	3401				CCGTCGGCAT	
	3451				CTGGGCATGG	
50	3501				CGACAGCATT	
50	3551				GCTATCTCAA	
	3601	TCCTACGGAC	GCTACAAAAA	CAGCATCAGC	CGCAGCACCG	GTGCGGACGA
	3651				GATGCAGCTG	
	3701				GAGATTTGAC	
	3751				GCATTCGCCG	
55	3801				TGAAGGCACG	
	3851	1 CGCGGGTCT	GAAGUTGTUG	CAACCETTGA	GCGATAAAGC	CGTCCTGTTT

	3901			CGACCTGAAC		
	3951			CTGCAGCAAC		
5	4001			GTTGCCGGCC		
3	4051			GGCACGTTAC		
	4101			GAGTCGGCGT	AGGCTACCGG	TTCCTCGAGC
	4151	ACCACCACCA	CCACCACTGA			
10						
10	1	MATNDDDVKK	AATVAIAAAY	NNGQEINGFK	AGETIYDIDE	DGTITKKDAT
	51			LTKTVNENKQ		
	101			LNKLGENITT		
	151			NTKADEAVKT	-	
	201			DKAEAVAAKV		
15	251			ATTEKLDTRL		
	301 351			LFQPYNVGGS CKDRSMLCAG		
	401			AIEAGYTGRG		
	451			KEAPEDGGGK		
	501			SVDGRPAGGI		
20	551			NSFGTTSRAG		
	601	YSGGDKTDEG	IRLMQQSDYG	NLSYHIRNKN	MLFIFSTGND	AQAQPNTYAL
	651			SGEKFKREMY		
	701			IAGTSFSAPI		
	751			GWGLLDAGKA		
25	801			GSQLQLHGNN		
	851 901			SLNSDGIVYL IIGGKLYMSA		
	951			ASLDSVEKTA		
	1001			NLENLMVELD		
	1051			NAADGVRIFN		
30	1101			QTQQDGGTWE		
	1151	TGENTTAAAT	LGMGRSTWSE	NSANAKTDSI	SLFAGIRHDA	GDIGYLKGLF
	1201			GSVNGTLMQL		
	1251	-		WSGNSLTEGT		
	1301			TGATAATGKT		VAGLGADVEF
35	1351	GNGWNGLARY	SYAGSKQYGN	HSGRVGVGYR	FLEHHHHHH*	
35						
40						
45						
50						
55						

	961cL-0	RF46.1				
	1		ТТССАТССАА	AGTACTGACC	ACAGCCATCC	TTGCCACTTT
	51			CCACAAACGA		
5	101			GCCTACAACA		
	151			CTACGACATT		
	201			CCGATGTTGA		
	251			ACTAACCTGA		
	301			AGTAAAAGCT		
	351			ACACTGATGC		
10	401			AACGCCTTGA		
	451			TAAGACAAAT		
	501			CCGTCGACAA		
	551			GAAACCAACA		
	601			ACAGACGGCC		
	651			CAGAAACTGC	-	
15	701			GCAGCCGACA		
	751			TGATATCGCT		
	801			ACGTGTACAC		
	851			CTGAACGCTA		
	901			ATCCATTGCC		
	951			CAGACCTGCG		
20	1001			TCCGGTCTGT		
				AGATTTGGCA		
	1051			TCGAACCCGA		
	1101					
	1151			GAGCGCAGCG		
25	1201			CAACCTGATG		
25	1251			TCCGCTTTTC		
	1301			GCCTCACATT		
	1351			CCTTTACCGC		
	1401	ACACCATCCC	GCCGACGGC'I'	ATGACGGGCC	ACAGGGCGGC	GGCTATCCCG
30						
	1451	CTCCCDDDGG	CGCGAGGGAT	ATATACAGCT	ΔCGΔCΔTΔΔΔ	AGGCGTTGCC
	1501			GACCGACAAC		
	1551			CCGGTAGTAT		
35	1601			CGATACAGCC		
	1651			CGGCACTGCA		
	1701			TCGGCGCAGG		
	1751			GTCATGCACG		
	1801			CAACGATTTG		
	1851			TCCGCGATTG		
40	1901			GTCAGCAATA		
				TCGGGGAAAA		
	1951					
	2001			CGCAGATGGG		
	2051			AATTTTGCCG		
	2101					AGCAGCGTTA
45	2151			CCTCAACCGT		
	2201			CGCCACCCGA		
	2251		TTCCGAATTT	TGAGAAGCAC	GTGAAATATG	ATACGTAACT
	2301	CGAG				

5 10	1 51 101 151 201 251 301 351 401 451 501 551 601 651 701 751	FKAGETIYDI KQNVDAKVKA TTFAEETKTN KTANEAKQTA KVTDIKADIA RLASAEKSIA GSGGGGSDLA IQSHQLGNLM SPVDGFSLYR QNIRLNLTDN NAAEAFNGTA ENKMARINDL IKGIGAVRGK	TAILATFCSG DEDGTITKKD AESEIEKLTT IVKIDEKLEA EETKQNVDAK TNKDNIAKKA DHDTRLNGLD NDSFIRQVLD IQQAAIKGNI IHWDGYEHHP RSTGQRLADR DIVKNIIGAA ADMAQLKDYA YGLGGITAHP SNLEQRYGKE VKYDT*	ATAADVEADD KLADTDAALA VADTVDKHAE VKAAETAAGK NSADVYTREE KTVSDLRKET RQHFEPDGKY GYIVRFSDHG ADGYDGPQGG FHNAGSMLTQ GEIVGAGDAV AAAIRDWAVQ IKRSQMGAIA	FKGLGLKKVV DTDAALDATT AFNDIADSLD AEAAAGTANT SDSKFVRIDG RQGLAEQAAL HLFGSRGELA HEVHSPFDNH GYPAPKGARD GVGDGFKRAT QGISEGSNIA NPNAAQGIEA LPKGKSAVSD	TNLTKTVNEN NALNKLGENI ETNTKADEAV AADKAEAVAA LNATTEKLDT SGLFQPYNVG ERSGHIGLGK ASHSDSDEAG IYSYDIKGVA RYSPELDRSG VMHGLGLLST VSNIFMAAIP NFADAAYAKY
20	961cL-7	41				
	1		TTCCATCCAA	AGTACTGACC	ACAGCCATCC	TTGCCACTTT
	51	CTGTAGCGGC	GCACTGGCAG	CCACAAACGA	CGACGATGTT	AAAAAAGCTG
	101	CCACTGTGGC	CATTGCTGCT	GCCTACAACA	ATGGCCAAGA	AATCAACGGT
	151	TTCAAAGCTG	GAGAGACCAT	CTACGACATT	GATGAAGACG	GCACAATTAC
	201	CAAAAAAGAC	GCAACTGCAG	CCGATGTTGA	AGCCGACGAC	TTTAAAGGTC
25	251	TGGGTCTGAA	AAAAGTCGTG	ACTAACCTGA	CCAAAACCGT	CAATGAAAAC
	301	AAACAAAACG	TCGATGCCAA	AGTAAAAGCT	GCAGAATCTG	AAATAGAAAA
	351	GTTAACAACC	AAGTTAGCAG	ACACTGATGC	CGCTTTAGCA	GATACTGATG
	401	CCGCTCTGGA	TGCAACCACC	AACGCCTTGA	ATAAATTGGG	AGAAAATATA
	451	ACGACATTTG	CTGAAGAGAC	TAAGACAAAT	ATCGTAAAAA	TTGATGAAAA
	501	ATTAGAAGCC	GTGGCTGATA	CCGTCGACAA	GCATGCCGAA	GCATTCAACG
30	551	ATATCGCCGA	TTCATTGGAT	GAAACCAACA	CTAAGGCAGA	CGAAGCCGTC
	601	AAAACCGCCA	ATGAAGCCAA	ACAGACGGCC	GAAGAAACCA	AACAAAACGT
	651	CGATGCCAAA	GTAAAAGCTG	CAGAAACTGC	AGCAGGCAAA	GCCGAAGCTG
	701	CCGCTGGCAC	AGCTAATACT	GCAGCCGACA	AGGCCGAAGC	TGTCGCTGCA
	751		ACATCAAAGC			
	801		AACAGTGCCG			
35	851		AATTGATGGT			
	901		CTGCTGAAAA			
	951		AAAACAGTGT			
	1001		AGCCGCGCTC			
	1051		GGGGTGGTGT			
40	1101		GCACCGCTCG			
40	1151		GTCCGTCAGG			
	1201		AAACTTATGG			
	1251		AAGGTCAGCC			
	1301		CATTACCTTG			
	1351		CCTTAACCGC			
45	1401		AAGATGGTTG			
·	1401	GCATTCCGGG	ANGAIGGIIG	CGAAACUCCA	OTTOMORNITO	GOCGNEATAG
	1451	CGGGCGAACA	TACATCTTTT	GACAAGCTTC	CCGAAGGCGG	CAGGGGGACA
50	1501		CGGCGTTCGG			
			TTCGCCGCCA			
	1551					
	1601		ACTCAATGTC			
	1651		ATGCCGTCAT			
	1701		TACTCCCTCG			
55	1751		GGAAGTGAAA			TATCGGCCTT
	1801	GCCGCCAAGC	AACTCGAGCA	CCACCACCAC	CACCACTGA	

	1	MKHFPSKVLT	TAILATFCSG	ALAATNDDDV	KKAATVAIAA	AYNNGOEING
	51				FKGLGLKKVV	
	101				DTDAALDATT	
		~				
5	151				AFNDIADSLD	
	201				AEAAAGTANT	
	251				SDSKFVRIDG	
	301	RLASAEKSIA	DHDTRLNGLD	KTVSDLRKET	RQGLAEQAAL	SGLFQPYNVG
	351	GSGGGGVAAD	IGAGLADALT	APLDHKDKGL	QSLTLDQSVR	KNEKLKLAAQ
	401	GAEKTYGNGD	SLNTGKLKND	KVSRFDFIRQ	IEVDGQLITL	ESGEFQVYKQ
10	451	SHSALTAFOT	EOTODSEHSG	KMVAKROFRI	GDIAGEHTSF	DKLPEGGRAT
10	501				EHLKSPELNV	
	551			_	QEVAGSAEVK	
	601	AAKQLEHHHH		1516116614	QD VNGDNL VN	1 414 G 1 1 1 1 1 1 1 1
	001	AANQLEIIIIIII	1111			
15						
15						
	0.61 at = 0	02				
	961cL-9		mmccamccaa	7 CM7 CMC7 CC	7 C7 CCC7 TCC	mmccca cmmm
	1				ACAGCCATCC	
	51				CGACGATGTT	
	101				ATGGCCAAGA	
20	151				GATGAAGACG	
	201	CAAAAAAGAC	GCAACTGCAG	CCGATGTTGA	AGCCGACGAC	TTTAAAGGTC
	251	TGGGTCTGAA	AAAAGTCGTG	ACTAACCTGA	CCAAAACCGT	CAATGAAAAC
	301	AAACAAAACG	TCGATGCCAA	AGTAAAAGCT	GCAGAATCTG	AAATAGAAAA
	351	GTTAACAACC	AAGTTAGCAG	ACACTGATGC	CGCTTTAGCA	GATACTGATG
	401	CCGCTCTGGA	TGCAACCACC	AACGCCTTGA	ATAAATTGGG	AGAAAATATA
25	451				ATCGTAAAAA	
	501				GCATGCCGAA	
	551				CTAAGGCAGA	
	601				GAAGAAACCA	
	651				AGCAGGCAAA	
30	701				AGGCCGAAGC	
	751				ACGAACAAAG	
	801				CAGAGAAGAG	
	851				CTACCGAAAA	
	901	CGCTTGGCTT	CTGCTGAAAA	ATCCATTGCC	GATCACGATA	CTCGCCTGAA
	951	CGGTTTGGAT	AAAACAGTGT	CAGACCTGCG	CAAAGAAACC	CGCCAAGGCC
35	1001	TTGCAGAACA	AGCCGCGCTC	TCCGGTCTGT	TCCAACCTTA	CAACGTGGGT
	1051	GGATCCGGCG	GAGGCGGCAC	TTCTGCGCCC	GACTTCAATG	CAGGCGGTAC
	1101	CGGTATCGGC	AGCAACAGCA	GAGCAACAAC	AGCGAAATCA	GCAGCAGTAT
	1151	CTTACGCCGG	TATCAAGAAC	GAAATGTGCA	AAGACAGAAG	CATGCTCTGT
	1201				AGGGATGCCA	
	1251	CCCCCCCCC	AATCTGCATA	CCGGAGACTT	TCCAAACCCA	AATGACGCAT
40	1301				TTGAAGCAGG	
40	1351				GGCGAATCCG	
	1401				ACACGGCTAT	
	1451				AAGCGCCTGA	
					GAGGCCGTTA	
	1501					
45	1551				AGAAATCGGA	
45	1601				TGGACGCAG	
	1651				ATGAATACGA	
	1701				CAATGCATGG	
	1751	GCGAACGTGG	CGTGCGCATC	GTCAATAACA	GTTTTGGAAC	AACATCGAGG
	1801	GCAGGCACTG	CCGACCTTTT	CCAAATAGCC	AATTCGGAGG	AGCAGTACCG
	1851	CCAAGCGTTG	CTCGACTATT	CCGGCGGTGA	TAAAACAGAC	GAGGGTATCC
50	1901				TGTCCTACCA	
	1951				AATGACGCAC	
	2001				AAAAGACGCT	
	2051				GAGAAAAGTT	
	2001	TIMICACAGI	CGCAGGCGIA	CACCGCAGIG	CHOMMAGII	CAMACGGGAA

	2101	ATGTATGGAG	AACCGGGTAC	AGAACCGCTT	GAGTATGGCT	CCAACCATTG
	2151	CGGAATTACT	GCCATGTGGT	GCCTGTCGGC	ACCCTATGAA	GCAAGCGTCC
	2201	GTTTCACCCG	TACAAACCCG	ATTCAAATTG	CCGGAACATC	CTTTTCCGCA
5	2251	CCCATCGTAA	CCGGCACGGC	GGCTCTGCTG	CTGCAGAAAT	ACCCGTGGAT
	2301	GAGCAACGAC	AACCTGCGTA	CCACGTTGCT	GACGACGGCT	CAGGACATCG
	2351	GTGCAGTCGG	CGTGGACAGC	AAGTTCGGCT	GGGGACTGCT	GGATGCGGGT
	2401	AAGGCCATGA	ACGGACCCGC	GTCCTTTCCG	TTCGGCGACT	TTACCGCCGA
	2451	TACGAAAGGT	ACATCCGATA	TTGCCTACTC	CTTCCGTAAC	GACATTTCAG
	2501	GCACGGGCGG	CCTGATCAAA	AAAGGCGGCA	GCCAACTGCA	ACTGCACGGC
10	2551	AACAACACCT	ATACGGGCAA	AACCATTATC	GAAGGCGGTT	CGCTGGTGTT
	2601	GTACGGCAAC	AACAAATCGG	ATATGCGCGT	CGAAACCAAA	GGTGCGCTGA
	2651	TTTATAACGG	GGCGGCATCC	GGCGGCAGCC	TGAACAGCGA	CGGCATTGTC
	2701	TATCTGGCAG	ATACCGACCA	ATCCGGCGCA	AACGAAACCG	TACACATCAA
	2751	AGGCAGTCTG	CAGCTGGACG	GCAAAGGTAC	GCTGTACACA	CGTTTGGGCA
4.5	2801	AACTGCTGAA	AGTGGACGGT	ACGGCGATTA	TCGGCGGCAA	GCTGTACATG
15	2851	TCGGCACGCG	GCAAGGGGGC	AGGCTATCTC	AACAGTACCG	GACGACGTGT
	2901	TCCCTTCCTG	AGTGCCGCCA	AAATCGGGCA	GGATTATTCT	TTCTTCACAA
	2951	ACATCGAAAC	CGACGGCGGC	CTGCTGGCTT	CCCTCGACAG	CGTCGAAAAA
	3001	ACAGCGGGCA	GTGAAGGCGA	CACGCTGTCC	TATTATGTCC	GTCGCGGCAA
	3051	TGCGGCACGG	ACTGCTTCGG	CAGCGGCACA	TTCCGCGCCC	GCCGGTCTGA
20	3101	AACACGCCGT	AGAACAGGGC	GGCAGCAATC	TGGAAAACCT	GATGGTCGAA
	3151		CCGAATCATC			
	3201		ACAGATATGC			
	3251	GCGCAGCGGC	AGCCGTACAG	CATGCGAATG	CCGCCGACGG	TGTACGCATC
	3301		TCGCCGCTAC			
	3351		GGACGCCGCC			
25	3401		TCTGCGCGTC			
	3451		GCGGTGTTGA			
	3501		GCGAAAACCG			
	3551		CAGCACATGG			
	3601		TGTTTGCAGG			
	3651		CTGTTCTCCT			
30	3701		GGACGAACAT			
	3751		CACTGGGCGG			
	3801		GAAGGCGGTC			
	3851		AGGCAGTGCT			
	3901		TCGGACTCGC			
35	3951		CTGTTTGCAA			
55	4001		GGTAACGGGC			
	4051		CACGCAATAT			
	4101		GAATTCGGCA			
	4151		CAAACAGTAC	GGCAACCACA	GCGGACGAGT	CGGCGTAGGC
	4201	TACCGGTTCT	GACTCGAG			

	1	MKHFPSKVLT	TAILATFCSG	ALAATNDDDV	KKAATVAIAA	AYNNGQEING
	51	FKAGETIYDI	DEDGTITKKD	ATAADVEADD	FKGLGLKKVV	TNLTKTVNEN
	101	KQNVDAKVKA	AESEIEKLTT	KLADTDAALA	DTDAALDATT	NALNKLGENI
_	151	TTFAEETKTN	IVKIDEKLEA	VADTVDKHAE	AFNDIADSLD	ETNTKADEAV
5	201	KTANEAKQTA	EETKQNVDAK	VKAAETAAGK	AEAAAGTANT	AADKAEAVAA
	251	KVTDIKADIA	TNKDNIAKKA	NSADVYTREE	SDSKFVRIDG	LNATTEKLDT
	301	RLASAEKSIA	DHDTRLNGLD	KTVSDLRKET	RQGLAEQAAL	SGLFQPYNVG
	351	GSGGGGTSAP	DFNAGGTGIG	SNSRATTAKS	AAVSYAGIKN	EMCKDRSMLC
	401	AGRDDVAVTD	RDAKINAPPP	NLHTGDFPNP	NDAYKNLINL	KPAIEAGYTG
10	451	RGVEVGIVDT	GESVGSISFP	ELYGRKEHGY	NENYKNYTAY	MRKEAPEDGG
10	501	GKDIEASFDD	EAVIETEAKP	TDIRHVKEIG	HIDLVSHIIG	GRSVDGRPAG
	551	GIAPDATLHI	MNTNDETKNE	MMVAAIRNAW	VKLGERGVRI	VNNSFGTTSR
	601	AGTADLFQIA	NSEEQYRQAL	LDYSGGDKTD	EGIRLMQQSD	YGNLSYHIRN
	651	KNMLFIFSTG	NDAQAQPNTY	ALLPFYEKDA	QKGIITVAGV	DRSGEKFKRE
	701	MYGEPGTEPL	EYGSNHCGIT	AMWCLSAPYE	ASVRFTRTNP	IQIAGTSFSA
15	751	PIVTGTAALL	LQKYPWMSND	NLRTTLLTTA	QDIGAVGVDS	KFGWGLLDAG
	801	KAMNGPASFP	FGDFTADTKG	TSDIAYSFRN	DISGTGGLIK	KGGSQLQLHG
	851	NNTYTGKTII	EGGSLVLYGN	NKSDMRVETK	GALIYNGAAS	GGSLNSDGIV
	901	YLADTDQSGA	NETVHIKGSL	QLDGKGTLYT	RLGKLLKVDG	TAIIGGKLYM
	951	SARGKGAGYL	NSTGRRVPFL	SAAKIGQDYS	FFTNIETDGG	LLASLDSVEK
	1001	TAGSEGDTLS	YYVRRGNAAR	TASAAAHSAP	AGLKHAVEQG	GSNLENLMVE
20	1051		ETVETAAADR			
	1101	FNSLAATVYA	DSTAAHADMQ	GRRLKAVSDG	LDHNGTGLRV	IAQTQQDGGT
0.5						
25	1151	WEQGGVEGKM	RGSTQTVGIA	AKTGENTTAA	ATLGMGRSTW	SENSANAKTD
	1201	SISLFAGIRH	DAGDIGYLKG	LFSYGRYKNS	ISRSTGADEH	AEGSVNGTLM
	1251	QLGALGGVNV	PFAATGDLTV	EGGLRYDLLK	QDAFAEKGSA	LGWSGNSLTE
	1301	GTLVGLAGLK	LSQPLSDKAV	LFATAGVERD	LNGRDYTVTG	GFTGATAATG
	1351	KTGARNMPHT	RLVAGLGADV	EFGNGWNGLA	RYSYAGSKQY	GNHSGRVGVG
30	1401	YRF*				
50						

**[0279]** It will be understood that the invention has been described by way of example only and modifications may be made whilst remaining within the scope and spirit of the invention. For instance, the use of proteins from other strains is envisaged [*e.g.* see WO00/66741 for polymorphic sequences for ORF4, ORF40, ORF46, 225, 235, 287, 519, 726, 919 and 953].

## **EXPERIMENTAL DETAILS**

35

40

45

50

55

## FPLC protein purification

[0280] The following table summarises the FPLC protein purification that was used:

Protein	PI	Column	Buffer	рН	Protocol
121.1 <sup>untagged</sup>	6.23	Mono Q	Tris	8.0	А
128.1untagged	5.04	Mono Q	Bis-Tris propane	6.5	А
406.1L	7.75	Mono Q	Diethanolamine	9.0	В
576.1L	5.63	Mono Q	Tris	7.5	В
593untagged	8.79	Mono S	Hepes	7.4	А
726 <sup>untagged</sup>	4.95	Hi-trap S	Bis-Tris	6.0	А
919untagged	10.5(-leader)	Mono S	Bicine	8.5	С
919Lorf4	10.4(-leader)	Mono S	Tris	8.0	В
920L	6.92(-leader)	Mono Q	Diethanolamine	8.5	А
953L	7.56(-leader)	Mono S	MES	6.6	D

(continued)

Protein	PI	Column	Buffer	рН	Protocol
982untagged	4.73	Mono Q	Bis-Tris propane	6.5	Α
919-287	6.58	Hi-trap Q	Tris	8.0	А
953-287	4.92	Mono Q	Bis-Tris propane	6.2	А

**[0281]** Buffer solutions included 20-120 mM NaCl, 5.0 mg/ml CHAPS and 10% v/v glycerol. The dialysate was centrifuged at 13000g for 20 min and applied to either a mono Q or mono S FPLC ion-exchange resin. Buffer and ion exchange resins were chosen according to the pl of the protein of interest and the recommendations of the FPLC protocol manual [Pharmacia: FPLC Ion Exchange and Chromatofocussing; Principles and Methods. Pharmacia Publication]. Proteins were eluted using a step-wise NaCl gradient. Purification was analysed by SDS-PAGE and protein concentration determined by the Bradford method.

**[0282]** The letter in the 'protocol' column refers to the following:

5

30

35

40

50

FPLC-A: Clones 121.1, 128.1, 593, 726, 982, periplasmic protein 920L and hybrid proteins 919-287, 953-287 were purified from the soluble fraction of E.coli obtained after disruption of the cells. Single colonies harbouring the plasmid of interest were grown overnight at 37°C in 20 ml of LB/Amp (100 μg/ml) liquid culture. Bacteria were diluted 1:30 in 1.0 L of fresh medium and grown at either 30°C or 37°C until the OD $_{550}$  reached 0.6-08. Expression of recombinant protein was induced with IPTG at a final concentration of 1.0 mM. After incubation for 3 hours, bacteria were harvested by centrifugation at 8000g for 15 minutes at 4°C. When necessary cells were stored at -20°C. All subsequent procedures were performed on ice or at 4°C. For cytosolic proteins (121.1, 128.1, 593, 726 and 982) and periplasmic protein 920L, bacteria were resuspended in 25 ml of PBS containing complete protease inhibitor (Boehringer-Mannheim). Cells were lysed by by sonication using a Branson Sonifier 450. Disrupted cells were centrifuged at 8000g for 30 min to sediment unbroken cells and inclusion bodies and the supernatant taken to 35% v/v saturation by the addition of 3.9 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The precipitate was sedimented at 8000g for 30 minutes. The supernatant was taken to 70% v/v saturation by the addition of 3.9 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and the precipitate collected as above. Pellets containing the protein of interest were identified by SDS-PAGE and dialysed against the appropriate ion-exchange buffer (see below) for 6 hours or overnight. The periplasmic fraction from E.coli expressing 953L was prepared according to the protocol of Evans et. al. [Infect.Immun. (1974), 10:1010-1017] and dialysed against the appropriate ion-exchange buffer. Buffer and ion exchange resin were chosen according to the pl of the protein of interest and the recommendations of the FPLC protocol manual (Pharmacia). Buffer solutions included 20 mM NaCl, and 10% (v/v) glycerol. The dialysate was centrifuged at 13000g for 20 min and applied to either a mono Q or mono S FPLC ion-exchange resin. Buffer and ion exchange resin were chosen according to the pl of the protein of interest and the recommendations of the FPLC protocol manual (Pharmacia). Proteins were eluted from the ion-exchange resin using either step-wise or continuous NaCl gradients. Purification was analysed by SDS-PAGE and protein concentration determined by Bradford method. Cleavage of the leader peptide of periplasmic proteins was demonstrated by sequencing the NH<sub>2</sub>-terminus (see below).

[0284] FPLC-B: These proteins were purified from the membrane fraction of E.coli. Single colonies harbouring the plasmid of interest were grown overnight at 37°C in 20 ml of LB/Amp (100 μg/ml) liquid culture. Bacteria were diluted 1:30 in 1.0 L of fresh medium. Clones 406.1L and 919LOrf4 were grown at 30°C and Orf25L and 576.1L at 37°C until the OD<sub>550</sub> reached 0.6-0.8. In the case of 919LOrf4, growth at 30°C was essential since expression of recombinant protein at 37°C resulted in lysis of the cells. Expression of recombinant protein was induced with IPTG at a final concentration of 1.0 mM. After incubation for 3 hours, bacteria were harvested by centrifugation at 8000g for 15 minutes at 4°C. When necessary cells were stored at -20 °C. All subsequent procedures were performed at 4°C. Bacteria were resuspended in 25 ml of PBS containing complete protease inhibitor (Boehringer-Mannheim) and lysed by osmotic shock with 2-3 passages through a French Press. Unbroken cells were removed by centrifugation at 5000g for 15 min and membranes precipitated by centrifugation at 100000g (Beckman Ti50, 38000rpm) for 45 minutes. A Dounce homogenizer was used to re-suspend the membrane pellet in 7.5 ml of 20 mM Tris-HCl (pH 8.0), 1.0 M NaCl and complete protease inhibitor. The suspension was mixed for 2-4 hours, centrifuged at 100000g for 45 min and the pellet resuspended in 7.5 ml of 20mM Tris-HCl (pH 8.0), 1.0M NaCl, 5.0mg/ml CHAPS, 10% (v/v) glycerol and complete protease inhibitor. The solution was mixed overnight, centrifuged at 100000g for 45 minutes and the supernatant dialysed for 6 hours against an appropriately selected buffer. In the case of Orf25.L, the pellet obtained after CHAPS extraction was found to contain the recombinant protein. This fraction, without further purification, was used to immunise mice.

**[0285] FPLC-C:** Identical to FPLC-A, but purification was from the soluble fraction obtained after permeabilising *E.coli* with polymyxin B, rather than after cell disruption.

[0286] FPLC-D: A single colony harbouring the plasmid of interest was grown overnight at 37°C in 20 ml of LB/Amp (100  $\mu$ g/ml) liquid culture. Bacteria were diluted 1:30 in 1.0 L of fresh medium and grown at 30°C until the OD<sub>550</sub> reached

0.6-0.8. Expression of recombinant protein was induced with IPTG at a final concentration of 1.0mM. After incubation for 3 hours, bacteria were harvested by centrifugation at 8000g for 15 minutes at 4°C. When necessary cells were stored at -20 °C. All subsequent procedures were performed on ice or at 4°C. Cells were resuspended in 20mM Bicine (pH 8.5), 20mM NaCl, 10% (v/v) glycerol, complete protease inhibitor (Boehringer-Mannheim) and disrupted using a Branson Sonifier 450. The sonicate was centrifuged at 8000g for 30 min to sediment unbroken cells and inclusion bodies. The recombinant protein was precipitated from solution between 35% v/v and 70% v/v saturation by the addition of 3.9M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The precipitate was sedimented at 8000g for 30 minutes, resuspended in 20 mM Bicine (pH 8.5), 20 mM NaCl, 10% (v/v) glycerol and dialysed against this buffer for 6 hours or overnight. The dialysate was centrifuged at 13000g for 20 min and applied to the FPLC resin. The protein was eluted from the column using a step-wise NaCl gradients. Purification was analysed by SDS-PAGE and protein concentration determined by Bradford method.

### Cloning strategy and oligonucleotide design

10

25

30

35

40

45

50

55

[0287] Genes coding for antigens of interest were amplified by PCR, using oligonucleotides designed on the basis of the genomic sequence of *N. meningitidis* B MC58. Genomic DNA from strain 2996 was always used as a template in PCR reactions, unless otherwise specified, and the amplified fragments were cloned in the expression vector pET21 b+ (Novagen) to express the protein as C-terminal His-tagged product, or in pET-24b+(Novagen) to express the protein in 'untagged' form (*e.g.*  $\Delta$ G 287K).

[0288] Where a protein was expressed without a fusion partner and with its own leader peptide (if present), amplification of the open reading frame (ATG to STOP codons) was performed.

[0289] Where a protein was expressed in 'untagged' form, the leader peptide was omitted by designing the 5'-end amplification primer downstream from the predicted leader sequence.

**[0290]** The melting temperature of the primers used in PCR depended on the number and type of hybridising nucleotides in the whole primer, and was determined using the formulae:

$$T_{m1} = 4 (G+C) + 2 (A+T)$$
 (tail excluded)

$$T_{m2} = 64.9 + 0.41 \text{ (% GC)} - 600/N$$
 (whole primer)

[0291] The melting temperatures of the selected oligonucleotides were usually 65-70°C for the whole oligo and 50-60°C for the hybridising region alone.

**[0292]** Oligonucleotides were synthesised using a Perkin Elmer 394 DNA/RNA Synthesizer, eluted from the columns in 2.0ml NH<sub>4</sub>OH, and deprotected by 5 hours incubation at 56°C. The oligos were precipitated by addition of 0.3M Na-Acetate and 2 volumes ethanol. The samples were centrifuged and the pellets resuspended in water.

		Sequences	Restricti site
OrfiL	Fwd	CGCGGATCCGCTAGC-AAAACAACCGACAAACGG	Nhel
	Rev	CCCGCTCGAG-TTACCAGCGGTAGCCTA	XhoI
Orf1	Fwd	CTAGCTAGC-GGACACACTTATTTCGGCATC	NheI
	Rev	CCCGCTCGAG- TTACCAGCGGTAGCCTAATTTG	Xhol
Orf1LOmpA	Fwd		Ndel-(Nhe
•	Rev	CCCGCTCGAG-	XhoI
Orf4L	Fwd	CGCGGATCCCATATG-AAAACCTTCTTCAAAACC	Ndel
	Rev	CCCGCTCGAG-TTATTTGGCTGCGCCTTC	XhoI
Orf7-1L	Fwd	GCGGCATTAAT-ATGTTGAGAAAATTGTTGAAAATGG	Asel
	Rev	GCGGCCTCGAG-TTATTTTTTCAAAATATATTTGC	Xhol
Orf9-1L	Fwd	GCGGCCATATG-TTACCTAACCGTTTCAAAATGT	Ndel
0119 12	Rev	GCGGCCTCGAG-TTATTTCCGAGGTTTTCGGG	XhoI
Orf23L	Fwd	CGCGGATCC <u>CATATG</u> -ACACGCTTCAAATATTC	NdeI
011202	Rev	CCCGCTCGAG-TTATTTAAACCGATAGGTAAA	XhoI
Orf25-1 His	Fwd	CGCGGATCCCATATG-GGCAGGGAAGAACCGC	NdeI
01123 1 1113	Rev	GCCCAAGCTT-ATCGATGGAATAGCCGCG	HindIII
Orf29-1 b-His	Fwd	CGCGGATCCGCTAGC-AACGGTTTGGATGCCCG	Nhel
(MC58)	Rev	CCCGCTCGAG-TTTGTCTAAGTTCCTGATAT	Xhol
(MC36)	Rev	CCCGCTCGAG-ATTCCCACCTGCCATC	, tho
Orf29-1 b-L	Fwd	CGCGGATCCGCTAGC-ATGAATTTGCCTATTCAAAAAT	NheI
(MC58)	Rev	CCCGCTCGAG-TTAATTCCCACCTGCCATC	Xhol
Orf29-1 c-His	Fwd	CGCGGATCCGCTAGC-ATGAATTTGCCTATTCAAAAAT	NheI
(MC58)	Rev	CCCGCTCGAG-TTGGACGATGCCCGCGA	XhoI
Orf29-1 c-L	Fwd	CGCGGATCCGCTAGC-ATGAATTTGCCTATTCAAAAAT	Nhel
(MC58)	Rev	CCCGCTCGAG-TTATTGGACGATGCCCGC	Xhol
Orf25L	Fwd	CGCGGATCCCATATG-TATCGCAAACTGATTGC	NdeI
	Rev	CCCGCTCGAG-CTAATCGATGGAATAGCC	XhoI
Orf37L	Fwd	CGCGGATCCCATATG-AAACAGACAGTCAAATG	NdeI
	Rev	CCCGCTCGAG-TCAATAACCCGCCTTCAG	XhoI
Orf38L	Fwd	CGCGGATCC <u>CATATG</u> - TTACGTTTGACTGCTTTAGCCGTATGCACC	Ndel
	Rev	CCCGCTCGAG- TTATTTTGCCGCGTTAAAAGCGTCGGCAAC	Xhol
Orf40L	Fwd	CGCGGATCCCATATG-AACAAAATATACCGCAT	Ndel
0111011	Rev	CCCGCTCGAG-TTACCACTGATAACCGAC	Xhol
Orf40.2-His	Fwd	CGCGGATCCCATATG-ACCGATGACGACGATTTAT	Ndel
J.110.8 1110	Rev	GCCCAAGCTT-CCACTGATAACCGACAGA	HindIII
Orf40.2L	Fwd	CGCGGATCCCATATG-AACAAAATATACCGCAT	Ndel
0.140.EL	Rev	GCCCAAGCTT-TTACCACTGATAACCGAC	HindIII
Orf46-2L	Fwd	GGGAATTCCATATG-GGCATTTCCCGCAAAATATC	Ndel
O1140-2L	Rev	CCCGCTCGAG-TTATTTACTCCTATAACGAGGTCTCTTAAC	XhoI
Orf46-2	Fwd	GGGAATTCCATATG-TCAGATTTGGCAAACGATTCTT	Ndel
O1140-2		CCCGCTCGAG-TTATTTACTCCTATAACGAGGTCTCTTAAC	Xhol
Orf46.1L	Rev	CCCGCTCGAG-TTATTTACTCCTATAACGAGGTCTCTTAAC	Ndel

	Rev	CCCGCTCGAG-TTACGTATCATATTTCACGTGC	Xhol
orf46. (His-GST)	Fwd	GGGAATTC <u>CATATG</u> CACGTGAAATATGATACGAAG	BamHI-NdeI
	Rev	CCCGCTCGAGTTTACTCCTATAACGAGGTCTCTTAAC	XhoI
orf46.1-His	Fwd	GGGAATTC <u>CATATG</u> TCAGATTTGGCAAACGATTCTT	Ndel
	Rev	CCCGCTCGAGCGTATCATATTTCACGTGC	Xhol
orf46.2-His	Fwd	GGGAATTC <u>CATATG</u> TCAGATTTGGCAAACGATTCTT	Ndel
	Rev	CCCGCTCGAGTTTACTCCTATAACGAGGTCTCTTAAC	Xhol
Orf65-1-(His/GST)	Fwd	CGCGGATCCCATATG-CAAAATGCGTTCAAAATCCC	BamHI-Ndel
(MC58)	Rev	CGCGGATCC <u>CATATG</u> -AACAAAATATACCGCAT	XhoI
		CCCGCTCGAG -TTTGCTTTCGATAGAACGG	
Orf72-1L	Fwd	GCGGC <u>CATATG</u> -GTCATAAAATATACAAATTTGAA	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TTAGCCTGAGACCTTTGCAAATT	XhoI
Orf76-1L	Fwd	GCGGC <u>CATATG</u> -AAACAGAAAAAAACCGCTG	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TTACGGTTTGACACCGTTTTC	XhoI
Orf83.1L	Fwd	CGCGGATCC <u>CATATG</u> -AAAACCCTGCTCCTC	Ndel
	Rev	CCCG <u>CTCGAG</u> -TTATCCTCCTTTGCGGC	Xhol
Orf85-2L	Fwd	GCGGC <u>CATATG</u> -GCAAAAATGATGAAATGGG	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TTATCGGCGCGGCGGCCC	Xhol
Orf91L (MC58)	Fwd	GCGGCCATATGAAAAAATCCTCCCTCATCA	Ndel
	Rev	GCGGCCTCGAGTTATTTGCCGCCGTTTTTGGC	Xhol
Orf91-His(MC58)	Fwd	GCGGCCATATGGCCCCTGCCGACGCGGTAAG	Ndel
	Rev	GCGGCCTCGAGTTTGCCGCCGTTTTTGGCTTTC	XhoI
Orf97-1L	Fwd	GCGGC <u>CATATG</u> -AAACACATACTCCCCCTGA	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TTATTCGCCTACGGTTTTTTG	Xhol
Orf119L (MC58)	Fwd	GCGGCCATATGATTTACATCGTACTGTTTC	NdeI
	Rev	GCGGCCTCGAGTTAGGAGAACAGGCGCAATGC	XhoI
Orf119-His(MC58)	Fwd	GCGGCCATATGTACAACATGTATCAGGAAAAC	Ndel
	Rev	GCGGCCTCGAGGGAGAACAGGCGCAATGCGG	Xhoi
Orf137.1 (His- GST) (MC58)	Fwd	CGCGGATCCGCTAGCTGCGGCACGGCGGG	BamHI-Nhel
, ( )	Rec	CCCGCTCGAGATAACGGTATGCCGCCAG	Xhol
Orf143-1L	Fwd	CGCGGATCC <u>CATATG</u> -GAATCAACACTTTCAC	Ndel
	Rev	CCCGCTCGAG-TTACACGCGGTTGCTGT	XhoI
008	Fwd	CGCGGATCCCATATG-AACAACAGACATTTTG	Ndel
	Rev	CCCGCTCGAG-TTACCTGTCCGGTAAAAG	XhoI
050-1(48)	Fwd	CGCGGATCCGCTAGC-ACCGTCATCAAACAGGAA	Nhel
• •	Rev	CCCGCTCGAG-TCAAGATTCGACGGGGA	Xhol
105	Fwd	CGCGGATCC <u>CATATG</u> -TCCGCAAACGAATACG	Ndel
	Rev	CCCGCTCGAG-TCAGTGTTCTGCCAGTTT	Xhol
111L	Fwd	CGCGGATCC <u>CATATG</u> -CCGTCTGAAACACG	Ndel
	Rev	CCCGCTCGAG-TTAGCGGAGCAGTTTTTC	XhoI
117-1	Fwd	CGCGGATCCCATATG-ACCGCCATCAGCC	Ndel
-	Rev	CCCGCTCGAG-TTAAAGCCGGGTAACGC	Xhol
121-1	Fwd	GCGGC <u>CATATG</u> -GAAACACAGCTTTACATCGG	Ndel
_	Rev	GCGGCCTCGAG-TCAATAATAATATCCCGCG	Xhol

122-1	Fwd	GCGGC <u>CATATG</u> -ATTAAAATCCGCAATATCC	NdeI
	Rev	GCGGC <u>CTCGAG</u> -TTAAATCTTGGTAGATTGGATTTGG	XhoI
128-1	Fwd	GCGGC <u>CATATG</u> -ACTGACAACGCACTGCTCC	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TCAGACCGCGTTGTCGAAAC	Xhol
148	Fwd	CGCGGATCC <u>CATATG</u> -GCGTTAAAAACATCAAA	Ndel
	Rev	CCCGCTCGAG-TCAGCCCTTCATACAGC	XhoI
149.1L (MC58)	Fwd	GCGGCATTAATGGCACAAACTACACTCAAACC	Asel
	Rev	GCGGCCTCGAGTTAAAACTTCACGTTCACGCCG	Xhol
149.1-His(MC58)	Fwd	GCGGCATTAATGCATGAAACTGAGCAATCGGTGG	Asel
	Rev	GCGGCCTCGAGAAACTTCACGTTCACGCCGCCGGTAAA	XhoI
205 (His-GST) (MC58)	Fwd	CGC <u>GGATCCCATATG</u> GGCAAATCCGAAAATACG	BamHI-Ndel
	Rev	CCCGCTCGAGATAATGGCGGCGGCGG	XhoI
206L	Fwd	CGCGGATCC <u>CATATG</u> -TTTCCCCCCGACAA	NdeI
	Rev	CCCG <u>CTCGAG</u> -TCATTCTGTAAAAAAAGTATG	Xhol
214 (His-GST) (MC58)	Fwd	CGC <u>GGATCCCATATG</u> CTTCAAAGCGACAGCAG	BamHI-Ndel
	Rev	CCCGCTCGAGTTCGGATTTTTGCGTACTC	XhoI
216	Fwd	CGCGGATCC <u>CATATG</u> -GCAATGGCAGAAAACG	NdeI
	Rev	CCCG <u>CTCGAG</u> -CTATACAATCCGTGCCG	XhoI
225-1L	Fwd	CGCGGATCC <u>CATATG</u> -GATTCTTTTTCAAACC	Ndel
	Rev	CCCG <u>CTCGAG</u> -TCAGTTCAGAAAGCGGG	Xhol
235L	Fwd	CGCGGATCC <u>CATATG</u> -AAACCTTTGATTTTAGG	Ndel
	Rev	CCCG <u>CTCGAG</u> -TTATTTGGGCTGCTCTTC	XhoI
243	Fwd	CGCGGATCC <u>CATATG</u> -GTAATCGTCTGGTTG	Ndel
	Rev	CCCG <u>CTCGAG</u> -CTACGACTTGGTTACCG	XhoI
247-1L	Fwd	GCGGC <u>CATATG</u> -AGACGTAAAATGCTAAAGCTAC	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TCAAAGTGTTCTGTTTGCGC	Xhol
264-His	Fwd	GCCGC <u>CATATG</u> -TTGACTTTAACCCGAAAAA	Ndel
	Rev	GCCGC <u>CTCGAG</u> -GCCGGCGGTCAATACCGCCCGAA	Xhoi
270 (His-GST) (MC58)	Fwd	CGC <u>GGATCCCATATG</u> GCGCAATGCGATTTGAC	BamHI-Ndel
	Rev	CCCGCTCGAGTTCGGCGGTAAATGCCG	XhoI
274L	Fwd	GCGGC <u>CATATG</u> -GCGGGGCCGATTTTTGT	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TTATTTGCTTTCAGTATTATTG	XhoI
283L	Fwd	GCGGC <u>CATATG</u> -AACTTTGCTTTATCCGTCA	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TTAACGGCAGTATTTGTTTAC	XhoI
285-His	Fwd	CGC <u>GGATCC</u> CATATGGGTTTGCGCTTCGGGC	BamHI
	Rev	GCCC <u>AAGCTT</u> TTTTCCTTTGCCGTTTCCG	HindIII
286-His	Fwd	CGCGGATCC <u>CATATG</u> -GCCGACCTTTCCGAAAA	Ndel
(MC58)	Rev	CCCGCTCGAG-GAAGCGCGTTCCCAAGC	XhoI
286L	Fwd	CGCGGATCC <u>CATATG</u> -CACGACACCCGTAC	Ndel
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTAGAAGCGCGTTCCCAA	XhoI
287L	Fwd	CTAGCTAGC-TTTAAACGCAGCGTAATCGCAATGG	NheI
	Rev	CCCG <u>CTCGAG</u> -TCAATCCTGCTCTTTTTTGCC	XhoI

	287	Fwd	CTAGCTAGC-GGGGGCGGCGGTGGCG	NheI
		Rev	CCCGCTCGAG-TCAATCCTGCTCTTTTTTGCC	Xhol
5	287LOrf4	Fwd	CTA <u>GCTAGC</u> GCTCATCCTCGCCGCC- TGCGGGGGCGGCGGT	NheI
		Rev	CCCG <u>CTCGAG</u> -TCAATCCTGCTCTTTTTTGCC	XhoI
	287-fu	Fwd	CGG <u>GGATCC</u> -GGGGGCGGCGGTGGCG	BamHl
		Rev	CCCG <u>CTCGAG</u> -TCAATCCTGCTCTTTTTTGCC	XhoI
10	287-His	Fwd	CTA <u>GCTAGC</u> -GGGGGGGGGGGGGGG	NheI
		Rev	CCCGCTCGAG-ATCCTGCTCTTTTTTGCC *	XhoI
	287-His(2996)	Fwd	CTAGCTAGC-TGCGGGGGGGGGGGGGGGGGGGGGGGGGGG	NheI
		Rev	CCCG <u>CTCGAG</u> -ATCCTGCTCTTTTTTGCC	XhoI
15	Δ1 287-His	Fwd	CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC §	NheI
	Δ2 287-His	Fwd	CGCGGATCCGCTAGC-CAAGATATGGCGGCAGT§	Nhel
	Δ3 287-His	Fwd	CGCGGATCC <u>GCTAGC</u> -GCCGAATCCGCAAATCA §	NheI
	Δ4 287-His	Fwd	CGCGCTAGC-GGAAGGGTTGATTTGGCTAATGG§	Nhel
20	Δ4 287MC58-His	Fwd	CGC <u>GCTAGC</u> -GGAAGGGTTGATTTGGCTAATGG§	Nhel
	287a-His	Fwd	CGC <u>CATATG</u> -TTTAAACGCAGCGTAATCGC	Ndel
		Rev	CCCG <u>CTCGAG</u> -AAAATTGCTACCGCCATTCGCAGG	XhoI
	287b-His	Fwd	CGC <u>CATATG</u> -GGAAGGGTTGATTTGGCTAATGG	Ndel
25	287b-2996-His	Rev	CCCGCTCGAG-CTTGTCTTTATAAATGATGACATATTTG	XhoI
	287b-MC58-His	Rev	CCCG <u>CTCGAG</u> -TTTATAAAAGATAATATATTGATTGATTCC	XhoI
	287c-2996-His	Fwd	CGC <u>GCTAGC</u> -ATGCCGCTGATTCCCGTCAATC §	Nhel
	'287 <sup>untagged</sup> ',(2996)	Fwd	CTAGCTAGC-GGGGGGGGGGGGGGG	Nhel
30		Rev	CCCG <u>CTCGAG</u> -TCAATCCTGCTCTTTTTTGCC	Xhol
	ΔG287-His *	Fwd	CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC	Nhel
		Rev	CCCG <u>CTCGAG</u> -ATCCTGCTCTTTTTTGCC	Xhol
	ΔG287K(2996)	Fwd	CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC	NheI
35		Rev	CCCG <u>CTCGAG</u> -TCAATCCTGCTCTTTTTTGCC	XhoI
	ΔG 287-L	Fwd	CGCGGATCC <u>GCTAGC</u> - TTTGAACGCAGTGTGATTGCAATGGCTTGTATTTTTGCC CTTTCAGCCTGT TCGCCCGATGTTAAATCGGCG	NheI
J		Rev	CCCGCTCGAG-TCAATCCTGCTCTTTTTTGCC	Xhol
40	ΔG 287-Orf4L	Fwd	CGCGGATCC <u>GCTAGC</u> - AAAACCTTCTTCAAAACCCTTTCCGCCGCCGCACTCGC CTCATCCTCGCCGCCTGC TCGCCCGATGTTAAATCG	Nhel
		Rev	CCCG <u>CTCGAG</u> -TCAATCCTGCTCTTTTTTGCC	Xhol
	292L	Fwd	CGCGGATCC <u>CATATG</u> -AAAACCAAGTTAATCAAA	Ndel
45		Rev	CCCG <u>CTCGAG</u> -TTATTGATTTTTGCGGATGA	Xhol
	308-1	Fwd	CGCGGATCC <u>CATATG</u> -TTAAATCGGGTATTTTATC	NdeI
		Rev	CCCGCTCGAG-TTAATCCGCCATTCCCTG	XhoI
	401L	Fwd	GCGGC <u>CATATG</u> -AAATTACAACAATTGGCTG	Ndel
50		Rev	GCGGC <u>CTCGAG</u> -TTACCTTACGTTTTTCAAAG	XhoI
	406L	Fwd	CGCGGATCC <u>CATATG</u> -CAAGCACGGCTGCT	Ndel
		Rev	CCCG <u>CTCGAG</u> -TCAAGGTTGTCCTTGTCTA	Xhol
	502-1L	Fwd	CGCGGATCC <u>CATATG</u> -ATGAAACCGCACAAC	Ndel
55		Rev	CCCG <u>CTCGAG</u> -TCAGTTGCTCAACACGTC	Xhol

502-A (His-GST)	Fwd	CGC <u>GGATCCCATATG</u> GTAGACGCGCTTAAGCA	BamHI-Ndel
	Rev	CCCG <u>CTCGAG</u> AGCTGCATGGCGGCG	XhoI
503-1L	Fwd	CGCGGATCC <u>CATATG</u> -GCACGGTCGTTATAC	Ndel
	Rev	CCCGCTCGAG-CTACCGCGCATTCCTG	XhoI
519-1L	Fwd	GCGGCCATATG-GAATTTTTCATTATCTTGTT	Ndel
	Rev	GCGGCCTCGAG-TTATTTGGCGGTTTTGCTGC	XhoI
525-1L	Fwd	GCGGC <u>CATATG</u> -AAGTATGTCCGGTTATTTTTC	Ndel
	Rev	GCGCCTCGAG-TTATCGGCTTGTGCAACGG	XhoI
529-(His/GST)	Fwd	CGCGGATCCGCTAGC-TCCGGCAGCAAAACCGA	Bam HI-NheI
(MC58)	Rev	GCCCAAGCTT-ACGCAGTTCGGAATGGAG	HindIII
552L	Fwd	GCCGCCATATGTTGAATATTAAACTGAAAACCTTG	Ndel
	Rev	GCCGCCTCGAGTTATTTCTGATGCCTTTTCCC	Xhol
556L	Fwd	GCCGCCATATGGACAATAAGACCAAACTG	Ndel
330L			XhoI
	Rev	GCCGCCTCGAGTTAACGGTGCGGACGTTTC	
557L	Fwd	CGCGGATCC <u>CATATG</u> -AACAAACTGTTTCTTAC	Ndel
	Rev	CCCGCTCGAG-TCATTCCGCCTTCAGAAA	XhoI
564ab-(His/GST) (MC58)	Fwd	CGC <u>GGATCCCATATG</u> - CAAGGTATCGTTGCCGACAAATCCGCACCT	BamHI-NdeI
	Rev	CCCG <u>CTCGAG</u> - AGCTAATTGTGCTTGGTTTGCAGATAGGAGTT	Xhol
564abL (MC58)	Fwd	CGCGGATCC <u>CATATG</u> - AACCGCACCCTGTACAAAGTTGTATTTAACAAACATC	NdeI
	Rev	CCCG <u>CTCGAG</u> - TTAAGCTAATTGTGCTTGGTTTGCAGATAGGAGTT	XhoI
564b- (His/GST)(MC58)	Fwd	CGCGGATCCCATATG- ACGGGAGAAAATCATGCGGTTTCACTTCATG	BamHI-Ndel
	Rev	CCCG <u>CTCGAG</u> - AGCTAATTGTGCTTGGTTTGCAGATAGGAGTT	Xhol
564c- (His/GST)(MC58)	Fwd	CGCGGATCCCATATG- GTTTCAGACGGCCTATACAACCAACATGGTGAAATT	BamHI-NdeI
(	Rev	CCCG <u>CTCGAG</u> - GCGGTAACTGCCGCTTGCACTGAATCCGTAA	XhoI
564bc- (His/GST)(MC58)	Fwd	CGCGGATCCCATATG- ACGGGAGAAAATCATGCGGTTTCACTTCATG	BamHI-NdeI
	Rev	CCCGCTCGAG- GCGGTAACTGCCGCTTGCACTGAATCCGTAA	XhoI
564d- (His/GST)(MC58)	Fwd	CGC <u>GGATCCCATATG</u> - CAAAGCAAAGTCAAAGCAGACCATGCCTCCGTAA	BamHI-Ndel
	Rev	CCCGCTCGAG- TCTTTTCCTTTCAATTATAACTTTAGTAGGTTCAATTTTG	XhoI
		GTCCCC	
564cd- (His/GST)(MC58)	Fwd		BamHI-NdeI
	Fwd	GTCCCC CGCGGATCCCATATG-	BamHI-NdeI XhoI
		GTCCCC  CGCGGATCCCATATG- GTTTCAGACGGCCTATACAACCAACATGGTGAAATT  CCCGCTCGAG- TCTTTTCCTTTCAATTATAACTTTAGTAGGTTCAATTTTG	
(His/GST)(MC58)	Rev	GTCCCC  CGCGGATCCCATATG- GTTTCAGACGGCCTATACAACCAACATGGTGAAATT  CCCGCTCGAG- TCTTTTCCTTTCAATTATAACTTTAGTAGGTTCAATTTTG GTCCCC	XhoI
(His/GST)(MC58)	Rev	GTCCCC  CGCGGATCCCATATG- GTTTCAGACGGCCTATACAACCAACATGGTGAAATT  CCCGCTCGAG- TCTTTTCCTTTCAATTATAACTTTAGTAGGTTCAATTTTG GTCCCC GCGGCCATATG-ACCCGTTTGACCCGCG	Xhol

580L	Fwd	GCGGC <u>CATATG</u> -GATTCGCCCAAGGTCGG	Ndel
	Rev	GCGGC <u>CTCGAG</u> -CTACACTTCCCCCGAAGTGG	XhoI
583L	Fwd	CGCGGATCC <u>CATATG</u> -ATAGTTGACCAAAGCC	Ndel
	Rev	CCCGCTCGAG-TTATTTTTCCGATTTTTCGG	XhoI
593	Fwd	GCGGC <u>CATATG</u> -CTTGAACTGAACGGACT	NdeI
	Rev	GCGGC <u>CTCGAG</u> -TCAGCGGAAGCGGACGATT	XhoI
650 (His-GST) (MC58)	Fwd	CGC <u>GGATCCCATATG</u> TCCAAACTCAAAACCATCG	BamHI-Ndel
ļ	Rev	CCCGCTCGAGGCTTCCAATCAGTTTGACC	Xhol
652	Fwd	GCGGC <u>CATATG</u> -AGCGCAATCGTTGATATTTTC	NdeI
į.	Rev	GCGGC <u>CTCGAG</u> -TTATTTGCCCAGTTGGTAGAATG	XhoI
664L	Fwd	GCGGC <u>CATATG</u> -GTGATACATCCGCACTACTTC	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TCAAAATCGAGTTTTACACCA	XhoI
726	Fwd	GCGGC <u>CATATG</u> -ACCATCTATTTCAAAAACGG	NdeI
	Rev	GCGGC <u>CTCGAG</u> -TCAGCCGATGTTTAGCGTCCATT	XhoI
741-His(MC58)	Fwd	CGCGGATCC <u>CATATG</u> -AGCAGCGGAGGGGGTG	NdeI
	Rev	CCCG <u>CTCGAG</u> -TTGCTTGGCGGCAAGGC	XhoI
ΔG741-His(MC58)	Fwd	CGCGGATCC <u>CATATG-</u> GTCGCCGCCGACATCG	Ndel
	Rev	CCCG <u>CTCGAG</u> -TTGCTTGGCGGCAAGGC	XhoI
686-2-(His/GST)	Fwd	CGC <u>GGATCCCATATG</u> -GGCGGTTCGGAAGGCG	BamHI-Nde
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTGAACACTGATGTCTTTTCCGA	Xhol
719-(His/GST)	Fwd	CGC <u>GGATCCGCTAGC</u> -AAACTGTCGTTGGTGTTAAC	BamHI-Nhe
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTGACCCGCTCCACGG	Xhol
730-His (MC58)	Fwd	GCCGCCATATGGCGGACTTGGCGCAAGACCC	Ndel
	Rev	GCCGCCTCGAGATCTCCTAAACCTGTTTTAACAATGCCG	XhoI
730A-His (MC58)	Fwd	GCCGCCATATGGCGGACTTGGCGCAAGACCC	Ndel
	Rev	GCGGCCTCGAGCTCCATGCTGTTGCCCCAGC	Xhol
730B-His (MC58)	Fwd	GCCGCCATATGGCGGACTTGGCGCAAGACCC	NdeI
	Rev	GCGGCCTCGAGAAAATCCCCGCTAACCGCAG	XhoI
741-His	Fwd	CGCGGATCC <u>CATATG</u> -AGCAGCGGAGGGGGTG	NdeI
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTGCTTGGCGGCAAGGC	XhoI
ΔG741-His	Fwd	CGCGGATCC <u>CATATG</u> -GTCGCCGCCGACATCG	NdeI
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTGCTTGGCGGCAAGGC	XhoI
743 (His-GST)	Fwd	CGC <u>GGATCCCATATG</u> GACGGTGTTGTGCCTGTT	BamHI-Nde
	Rev	CCCGCTCGAGCTTACGGATCAAATTGACG	XhoI
757 (His-GST) (MC58)	Fwd	CGC <u>GGATCCCATATG</u> GGCAGCCAATCTGAAGAA	BamHI-Nde
	Rev	CCCGCTCGAGCTCAGCTTTTGCCGTCAA	XhoI
759-His/GST	Fwd	CGCGGATCCGCTAGC-TACTCATCCATTGTCCGC	BamHI-Nhe
(MC58)	Rev	CCCG <u>CTCGAG</u> -CCAGTTGTAGCCTATTTTG	XhoI
759L	Fwd	CGCGGATCC <u>GCTAGC</u> -ATGCGCTTCACACACAC	Nhel
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTACCAGTTGTAGCCTATTT	Xhol
760-His	Fwd	GCCGCCATATGGCACAAACGGAAGGTTTGGAA	NdeI
	Rev	GCCGCCTCGAGAAAACTGTAACGCAGGTTTGCCGTC	Xhol
769-His (MC58)	Fwd	GCGGCCATATGGAAGAAACACCGCGCGAACCG	Ndel

	Rev	GCGGCCTCGAGGAACGTTTTATTAAACTCGAC	Xhol
907L	Fwd	GCGGC <u>CATATG</u> -AGAAAACCGACCGATACCCTA	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TCAACGCCACTGCCAGCGGTTG	Xhol
911L	Fwd	CGCGGATCC <u>CATATG</u> -AAGAAGAACATATTGGAATTTTGGGTCGGACTG	Ndel
	Rev	CCCGCTCGAG-TTATTCGGCGGCTTTTTCCGCATTGCCG	Xhol
911LOmpA	Fwd	GGGAATTC <u>CATATG</u> AAAAAGACAGCTATCGCGATTGCA GTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCC <u>GC</u> TAGC-GCTTTCCGCGTGGCCGGCGGTGC	Ndel-(Nhel
	Rev	CCCG <u>CTCGAG</u> -TTATTCGGCGGCTTTTTCCGCATTGCCG	XhoI
911LPelB	Fwd	CATG <u>CCATGG</u> -CTTTCCGCGTGGCCGGCGGTGC	Ncol
	Rev	CCCGCTCGAG-TTATTCGGCGGCTTTTTCCGCATTGCCG	XhoI
913-His/GST	Fwd	CGCGGATCCCATATG-TTTGCCGAAACCCGCC	BamHI-Nde
(MC58)	Rev	CCCG <u>CTCGAG</u> -AGGTTGTTCCAGGTTG	XhoI
913L	Fwd	CGCGGATCC <u>CATATG</u> -AAAAAAACCGCCTATG	Ndel
(MC58)	Rev	CCCGCTCGAG-TTAAGGTTGTGTTCCAGG	XhoI
919L	Fwd	CGCGGATCC <u>CATATG</u> -AAAAAATACCTATTCCGC	Ndel
	Rev	CCCG <u>CTCGAG</u> -TTACGGGCGGTATTCGG	XhoI
919	Fwd	CGCGGATCC <u>CATATG</u> -CAAAGCAAGAGCATCCAAA	Ndel
	Rev	CCCGCTCGAG-TTACGGGCGGTATTCGG	XhoI
919L Orf4	Fwd	GGGAATTC <u>CATATG</u> AAAACCTTCTTCAAAACCCTTTCCG CCGCCGC <u>GCTAGC</u> GCTCATCCTCGCCGCC- TGCCAAAGCAAGAGCATC	NdeI-(NheI
	Rev	CCCGCTCGAG-TTACGGGCGGTATTCGGGCTTCATACCG	Xhol
(919)-287fusion	Fwd	CGCGGATCCGTCGAC-TGTGGGGGCGGCGGTGGC	SalI
	Rev	CCCGCTCGAG-TCAATCCTGCTCTTTTTTGCC	Xhol
920-1L	Fwd	GCGGC <u>CATATG</u> -AAGAAAACATTGACACTGC	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TTAATGGTGCGAATGACCGAT	Xhol
925-His/GST (MC58) GATE	Fwd	ggggacaagtttgtacaaaaaagcaggctTGCGGCAAGGATGCCGG	attB1
	Rev	ggggaccactttgtacaagaaagctgggtCTAAAGCAACAATGCCGG	attB2
926L	Fwd	CGCGGATCC <u>CATATG</u> -AAACACACCGTATCC	NdeI
	Rev	CCCG <u>CTCGAG</u> -TTATCTCGTGCGCGCC	XhoI
927-2-(His/GST)	Fwd	CGCGGATCCCATATG-AGCCCCGCGCCGATT	BamHI-Nd
(MC58)	Rev	CCCGCTCGAG-TTTTTGTGCGGTCAGGCG	Xhol
932-His/GST (MC58) GATE	Fwd	ggggacaagtttgtacaaaaaagcaggctTGTTCGTTTGGGGGATTTAA ACCAAACCAAATC	attB1
935 (His-GST) (MC58)	For	CGC <u>GGATCCCATATG</u> GCGGATGCGCCCGCG	BamHI-Nde
	Rev	CCCG <u>CTCGAG</u> AAACCGCCAATCCGCC	Xhol
	Rev	ggggaccactttgtacaagaaagctgggtTCATTTTGTTTTTCCTTCTTCTCGAGGCCATT	attB2
936-1L	Fwd	CGCGGATCC <u>CATATG</u> -AAACCCAAACCGCAC	Ndel
	Rev	CCCG <u>CTCGAG</u> -TCAGCGTTGGACGTAGT	Xhol
953L	Fwd	GGGAATTC <u>CATATG</u> -AAAAAAATCATCTTCGCCG	Ndel
	Rev	CCCGCTCGAG-TTATTGTTTGGCTGCCTCGAT	Xhol
953-fu	Fwd	GGGAATTCCATATG-GCCACCTACAAAGTGGACG	Ndel
	Rev	CGGGGATCC-TTGTTTGGCTGCCTCGATTTG	BamHI

954 (His-GST) (MC58)	Fwd	CGC <u>GGATCCCATATG</u> CAAGAACAATCGCAGAAAG	BamHI-NdeI
, ,	Rev	CCCGCTCGAGTTTTTTCGGCAAATTGGCTT	XhoI
958-His/GST (MC58) GATE	Fwd	ggggacaagtttgtacaaaaaagcaggctGCCGATGCCGTTGCGG	attB1
	Rev	ggggaccactttgtacaagaaagctgggtTCAGGGTCGTTTGTTGCG	attB2
961L	Fwd	CGCGGATCC <u>CATATG</u> -AAACACTTTCCATCC	NdeI
_	Rev	CCCG <u>CTCGAG</u> -TTACCACTCGTAATTGAC	XhoI
961	Fwd	CGCGGATCC <u>CATATG</u> -GCCACAAGCGACGAC	Ndel
	Rev	CCCG <u>CTCGAG</u> -TTACCACTCGTAATTGAC	XhoI
961 c (His/GST)	Fwd	CGCGGATCCCATATG-GCCACAAACGACG	BamHI-NdeI
	Rev	CCCG <u>CTCGAG</u> -ACCCACGTTGTAAGGTTG	XhoI
961 c-(His/GST)	Fwd	CGCGGATCCCATATG-GCCACAAGCGACGACGA	BamHI-NdeI
(MC58)	Rev	CCCGCTCGAG-ACCCACGTTGTAAGGTTG	XhoI
961 c-L	Fwd	CGCGGATCC <u>CATATG</u> -ATGAAACACTTTCCATCC	Ndel
	Rev	CCCGCTCGAG-TTAACCCACGTTGTAAGGT	XhoI
961 c-L	Fwd	CGCGGATCC <u>CATATG</u> -ATGAAACACTTTCCATCC	NdeI
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTAACCCACGTTGTAAGGT	XhoI
961 d (His/GST)	Fwd	CGCGGATCCCATATG-GCCACAAACGACG	BamHI-Ndel
	Rev	CCCGCTCGAG-GTCTGACACTGTTTTATCC	XhoI
961 Δ1-L	Fwd	CGCGGATCCCATATG-ATGAAACACTTTCCATCC	NdeI
	Rev	CCCGCTCGAG-TTATGCTTTGGCGGCAAAG	XhoI
fu 961	Fwd	CGCGGATCCCATATG- GCCACAAACGACGAC	Ndel
	Rev	CGCGGATCC-CCACTCGTAATTGACGCC	BamHI
fu 961	Fwd	CGCGGATCC <u>CATATG</u> -GCCACAAGCGACGAC	Ndel
(MC58)	Rev	CGCGGATCC-CCACTCGTAATTGACGCC	BamHI
fu 961 c	Fwd	CGCGGATCC <u>CATATG</u> -GCCACAAACGACGAC	Ndel
	Rev	CGCGGATCC -ACCCACGTTGTAAGGTTG	BamHI
fu 961 c-L	Fwd	CGCGGATCCCATATG- ATGAAACACTTTCCATCC	Ndel
	Rev	CGCGGATCC -ACCCACGTTGTAAGGTTG	BamHI
fu (961 )-	Fwd	CGCGGATCC -GGAGGGGGTGTCG	BamHI
741(MC58)-His			
707	Rev	CCCG <u>CTCGAG</u> -TTGCTTGGCGGCAAGGC	XhoI
fu (961 )-983-His	Fwd	CGC <u>GGATCC</u> - GGCGGAGGCGCACTT	BamHI
	Rev	CCCG <u>CTCGAG</u> -GAACCGGTAGCCTACG	XhoI
fu (961)- Orf46.1-	Fwd	CGCGGATCCGGTGGTGGT-	BamHI
His	Dave	TCAGATTTGGCAAACGATTC	Vhal
0 (0(1 1)	Rev	CCCGCTCGAG-CGTATCATATTTCACGTGC	Xhol
fu (961 c-L)- 741(MC58)	Fwd	CGCGGATCC -GGAGGGGGTGGTGTCG	BamHI
	Rev	CCCG <u>CTCGAG</u> -TTATTGCTTGGCGGCAAG	XhoI
fu (961c-L )-983	Fwd	CGCGGATCC - GGCGGAGGCGGCACTT	BamHI
	Rev	CCCGCTCGAG-TCAGAACCGGTAGCCTAC	Xhol
fu (961c-L)- Orf46.1	Fwd	CGC <u>GGATCC</u> GGTGGTGGT- TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCCGCTCGAG-TTACGTATCATATTTCACGTGC	XhoI
061-(His/GST)	Fwd	CGCGGATCCCATATG-GCCACAAGCGACGACG	BamHI-Ndel

(MC58)	Rev	CCCG <u>CTCGAG</u> -CCACTCGTAATTGACGCC	XhoI
961 Δ1-His	Fwd	CGCGGATCC <u>CATATG</u> -GCCACAAACGACGAC	NdeI
	Rev	CCCG <u>CTCGAG</u> -TGCTTTGGCGGCAAAGTT	XhoI
961a-(His/GST)	Fwd	CGCGGATCC <u>CATATG</u> -GCCACAAACGACGAC	BamHI-Nde
	Rev	CCCG <u>CTCGAG</u> -TTTAGCAATATTATCTTTGTTCGTAGC	XhoI
961b-(His/GST)	Fwd	CGCGGATCC <u>CATATG</u> -AAAGCAAACCGTGCCGA	BamHI-Nde
	Rev	CCCG <u>CTCGAG</u> -CCACTCGTAATTGACGCC	XhoI
961-His/GST GATE	Fwd	ggggacaagtttgtacaaaaaagcaggctGCAGCCACAAACGACGACGATGTTAAAAAAGC	attB1
	Rev	ggggaccactttgtacaagaaagctgggtTTACCACTCGTAATTGACGC CGACATGGTAGG	attB2
982	Fwd	GCGGC <u>CATATG</u> -GCAGCAAAAGACGTACAGTT	Ndel
	Rev	GCGGC <u>CTCGAG</u> -TTACATCATGCCGCCCATACCA	XhoI
983-His (2996)	Fwd	CGCGGATCC <u>GCTAGC</u> -TTAGGCGGCGGCGGAG	Nhel
	Rev	CCCG <u>CTCGAG</u> -GAACCGGTAGCCTACG	XhoI
ΔG983-His (2996)	Fwd	CCCCTAGCTAGC-ACTTCTGCGCCCGACTT	Nhel
	Rev	CCCG <u>CTCGAG</u> -GAACCGGTAGCCTACG	Xhol
983-His	Fwd	CGCGGATCC <u>GCTAGC</u> -TTAGGCGGCGGCGGAG	Nhel
	Rev	CCCG <u>CTCGAG</u> -GAACCGGTAGCCTACG	XhoI
ΔG983-His	Fwd	CGCGGATCC <u>GCTAGC</u> -ACTTCTGCGCCCGACTT	Nhel
	Rev	CCCG <u>CTCGAG</u> -GAACCGGTAGCCTACG	XhoI
983L	Fwd	CGCGGATCCGCTAGC- CGAACGACCCCAACCTTCCCTACAAAAACTTTCAA	Nhel
	Rev	CCCG <u>CTCGAG</u> -TCAGAACCGACGTGCCAAGCCGTTC	XhoI
987-His (MC58)	Fwd	GCCGCCATATGCCCCCACTGGAAGAACGGACG	NdeI
	Rev	GCCGCCTCGAGTAATAAACCTTCTATGGGCAGCAG	XhoI
989-(His/GST)	Fwd	CGCGGATCCCATATG-TCCGTCCACGCATCCG	BamHI-Nde
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTTGAATTTGTAGGTGTATTG	Xhol
989L	Fwd	CGCGGATCC <u>CATATG</u> -ACCCCTTCCGCACT	Ndel
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTATTTGAATTTGTAGGTGTAT	XhoI
CrgA-His	Fwd	CGCGGATCC <u>CATATG</u> -AAAACCAATTCAGAAGAA	NdeI
(MC58)	Rev	CCCG <u>CTCGAG</u> -TCCACAGAGATTGTTTCC	XhoI
PilC1-ES	Fwd	GATGCCCGAAGGCCGG	
(MC58)	Rev	GCCCAAGCTT-TCAGAAGAAGACTTCACGC	
PilC1-His	Fwd	CGCGGATCC <u>CATATG</u> -CAAACCCATAAATACGCTATT	NdeI
(MC58)	Rev	GCCC <u>AAGCTT</u> -GAAGAAGACTTCACGCCAG	HindIII
Δ1PilC1-His	Fwd	CGCGGATCC <u>CATATG</u> -GTCTTTTTCGACAATACCGA	Ndel
(MC58)	Rev	GCCC <u>AAGCTT</u> -	HindllI
PilC1L	Fwd	CGCGGATCCCATATG-AATAAAACTTTAAAAAGGCGG	Ndel
(MC58)	Rev	GCCC <u>AAGCTT</u> -TCAGAAGAAGACTTCACGC	HindIII
ΔGTbp2-His	Fwd	CGCGAATCC <u>CATATG</u> -TTCGATCTTGATTCTGTCGA	Ndel
(MC58)	Rev	CCCG <u>CTCGAG</u> -TCGCACAGGCTGTTGGCG	XhoI
Tbp2-His	Fwd	CGCGAATCC <u>CATATG</u> -TTGGGCGGAGGCGGCAG	Ndel
(MC58)	Rev	CCCGCTCGAG-TCGCACAGGCTGTTGGCG	Xhol
Tbp2-His(MC58)	Fwd	CGCGAATCC <u>CATATG</u> -TTGGGCGGAGGCGGCAG	Ndel
-	Rev	CCCGCTCGAG-TCGCACAGGCTGTTGGCG	XhoI

NMB0109- (His/GST)	Fwd	CGC <u>GGATCCCATATG</u> -GCAAATTTGGAGGTGCGC	BamHI-Nde
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTCGGAGCGGTTGAAGC	Xhol
NMB0109L	Fwd	CGCGGATCC <u>CATATG</u> -CAACGTCGTATTATAACCC	Ndel
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTATTCGGAGCGGTTGAAG	XhoI
NMB0207- (His/GST)	Fwd	CGC <u>GGATCCCATATG</u> - GGCATCAAAGTCGCCATCAACGGCTAC	BamHI-Nd
(MC58)	Rev	CCCG <u>CTCGAG</u> -TTTGAGCGGGCGCACTTCAAGTCCG	XhoI
NMB0462- (His/GST)	Fwd	CGC <u>GGATCCCATATG</u> -GGCGGCAGCGAAAAAAAC	BamHI-Nd
(MC58)	Rev	CCCG <u>CTCGAG</u> -GTTGGTGCCGACTTTGAT	XhoI
NMB0623- (His/GST)	Fwd	CGC <u>GGATCCCATATG</u> -GGCGGCGGAAGCGATA	BamHI-Nd
(MC58)	Rev	CCCGCTCGAG-TTTGCCCGCTTTGAGCC	Xhol
NMB0625 (His-GST)(MC58)	Fwd	CGC <u>GGATCCCATATG</u> GGCAAATCCGAAAATACG	BamHI-Nd
<u> </u>	Rev	CCCGCTCGAGCATCCCGTACTGTTTCG	XhoI
NMB0634 (His/GST)(MC58)	Fwd	ggggacaagtttgtacaaaaaagcaggctCCGACATTACCGTGTACAAC GGCCAACAAGAA	attB1
:	Rev	ggggaccactttgtacaagaaagctgggtCTTATTTCATACCGGCTTGCT CAAGCAGCCGG	attB2
NMB0776- His/GST (MC58)	Fwd	ggggacaagtttgtacaaaaagcaggctGATACGGTGTTTTCCTGTAA AACGGACAACAA	attB1
GA16	Rev	ggggaccactttgtacaagaaagctgggtCTAGGAAAAATCGTCATCGT TGAAATTCGCC	attB2
NMB1115-		ggggacaagtttgtacaaaaaagcaggctATGCACCCCATCGAAACC	attB1
His/GST (MC58)	Rev	ggggaccactttgtacaagaaagctgggtCTAGTCTTGCAGTGCCTC	atiB2
NMB1343- (His/GST)	Fwd	CGC <u>GGATCCCATATG</u> - GGAAATTTCTTATATAGAGGCATTAG	BamHI-Nd
(MC58)	Rev	CCCG <u>CTCGAG</u> - GTTAATTTCTATCAACTCTTTAGCAATAAT	XhoI
NMB1369 (His- GST (MC58)	Fwd	CGC <u>GGATCCCATATG</u> GCCTGCCAAGACGACA	BamHI-Nd
	Rev	CCCGCTCGAGCCGCCTCCTGCCGAAA	XhoI
NMB1551 (His- GST)(MC58)	Fwd	CGC <u>GGATCCCATATG</u> GCAGAGATCTGTTTGATAA	BamHI-Nd
	Rev	CCCGCTCGAGCGGTTTTCCGCCCAATG	XhoI
NMB1899 (His- GST) (MC58)	Fwd	CGC <u>GGATCCCATATG</u> CAGCCGGATACGGTC	BamHI-Nd
	Rev	CCCGCTCGAGAATCACTTCCAACACAAAAT	Xhol
NMB2050- (His/GST)	Fwd	CGC <u>GGATCCCATATG</u> -TGGTTGCTGATGAAGGGC	BamHI-Nd
(MC58)	Rev	CCCG <u>CTCGAG</u> -GACTGCTTCATCTTCTGC	XhoI
NMB2050L	Fwd	CGCGGATCC <u>CATATG</u> -GAACTGATGACTGTTTTGC	Ndel
(MC58)	Rev	CCCGCTCGAG-TCAGACTGCTTCATCTTCT	Xhol
NMB2159- (His/GST)	Fwd	CGCGGATCCCATATG- AGCATTAAAGTAGCGATTAACGGTTTCGGC	BamHI-Nd
(MC58)	Rev	CCCGCTCGAG- GATTTTGCCTGCGAAGTATTCCAAAGTGCG	Xhol
fu-∆G287His	Fwd	CGCGGATCC <u>GCTAGC</u> -CCCGATGTTAAATCGGC	Nhel

	Rev	CGG <u>GGATCC</u> -ATCCTGCTCTTTTTTGCCGG	BamHI
fu-(ΔG287)-919-	Fwd	CGC <u>GGATCC</u> GGTGGTGGT-	BamHI
His		CAAAGCAAGAGCATCCAAACC	
	Rev	CCC <u>AAGCTT</u> -TTCGGGCGGTATTCGGGCTTC	HindIII
fu-(ΔG287)-953- His	Fwd	CGC <u>GGATCC</u> GGTGGTGGT- GCCACCTACAAAGTGGAC	BamHI
	Rev	GCCC <u>AAGCTT</u> -TTGTTTGGCTGCCTCGAT	HindIII
fu-(ΔG287)-961-	Fwd	CGC <u>GGATCC</u> GGTGGTGGTGGT-ACAAGCGACGACG	BamHI
His	Rev	GCCC <u>AAGCTT</u> -CCACTCGTAATTGACGCC	HindIII
fu-(ΔG287)- Orf46.1-His	Fwd	CGC <u>GGATCC</u> GGTGGTGGT- TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCC <u>AAGCTT</u> -CGTATCATATTTCACGTGC	HindIII
fu-(ΔG287-919)- Orf46.1-His	Fwd	CCC <u>AAGCTT</u> GGTGGTGGTGGT- TCAGATTTGGCAAACGATTC	HindIII
	Rev	CCC <u>GCTCGAG</u> -CGTATCATATTTCACGTGC	Xhol
fu-(ΔG287- Orf46.1)-919-His	Fwd	CCC <u>AAGCTT</u> GGTGGTGGTGGT- CAAAGCAAGAGCATCCAAACC	HindIII
,	Rev	CCCGCTCGAG-CGGGCGGTATTCGGGCTT	Xhol
fu ΔG287(394.98)- 	Fwd	CGCGGATCC <u>GCTAGC</u> -CCCGATGTTAAATCGGC	Nhel
	Rev	CGG <u>GGATCC</u> -ATCCTGCTCTTTTTTGCCGG	BamHI
fu Orf1-(Orf46.1)-	Fwd	CGCGGATCCGCTAGC-GGACACACTTATTTCGGCATC	NheI
His	Rev	CGCGGATCC-CCAGCGGTAGCCTAATTTGAT	
fu (Orf1)-Orf46.1- His	Fwd	CGC <u>GGATCC</u> GGTGGTGGT- TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCC <u>AAGCTT</u> -CGTATCATATTTCACGTGC	HindIII
fu (919)-Orf46.1-	Fwd1	GCGGC <u>GTCGAC</u> GGTGGCGGAGGCACTGGATCCTCAG	Sall
His	Fwd2	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC	
	Rev	CCC <u>GCTCGAG</u> -CGTATCATATTTCACGTGC	Xhol
Fu orf46	Fwd	GGAATTC <u>CATATG</u> TCAGATTTGGCAAACGATTC	Ndel
	Rev	CGC <u>GGATCC</u> CGTATCATATTTCACGTGC	BamHI
Fu (orf46)-287-His	Fwd	CGG <u>GGATCC</u> GGGGGGGGGGGGGGG	BamHI
	Rev	CCC <u>AAGCTT</u> ATCCTGCTCTTTTTTGCCGGC	HindIII
Fu (orf46)-919-His	Fwd	CGC <u>GGATCC</u> GGTGGTGGTGGTCAAAGCAAGAGCATCCA AACC	BamHI
	Rev	CCC <u>AAGCTT</u> CGGGCGGTATTCGGGCTTC	HindIII
Fu (orf46-919)- 287-His	Fwd	CCCC <u>AAGCTT</u> GGGGGCGCGGTGGCG	HindIII
	-	CCCGCTCGAGATCCTGCTCTTTTTTGCCGGC	Xhol
	Rev	<u> </u>	
Fu (orf46-287)- 919-His	Fwd	CCC <u>AAGCTT</u> GGTGGTGGTGGTCAAAGCAAGAGCAT CCAAACC	HindIII
	-	CCC <u>AAGCTT</u> GGTGGTGGTGGTCAAAGCAAGAGCAT	HindIII XhoI
	Fwd Rev Fwd1	CCCAAGCTTGGTGGTGGTGGTCAAAGCAAGAGCAT CCAAACC CCCGCTCGAGCGGGCGGTATTCGGGCTT GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	
919-His	Fwd Rev Fwd1	CCCAAGCTTGGTGGTGGTGGTCAAAGCAAGAGCAT CCAAACC CCCGCTCGAGCGGGCGGTATTCGGGCTT GGAGGCACTGGATCCGCAGCCACAAACGACGACGA GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	XhoI XhoI
919-His	Fwd Rev Fwd1	CCCAAGCTTGGTGGTGGTGGTCAAAGCAAGAGCAT CCAAACC CCCGCTCGAGCGGGCGGTATTCGGGCTT GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
919-His	Fwd Rev Fwd1 Fwd2	CCCAAGCTTGGTGGTGGTGGTCAAAGCAAGAGCAT CCAAACC CCCGCTCGAGCGGGCGGTATTCGGGCTT GGAGGCACTGGATCCGCAGCCACAAACGACGACGA GCGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG CCCGCTCGAG-ACCCAGCTTGTAAGGTTG GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI XhoI

(ΔG741 )-983-His	Fwd	GCGGC <u>CTCGAG</u> -	XhoI
		GGATCCGGCGGAGGCGCACTTCTGCG	
	Rev	CCCG <u>CTCGAG</u> -GAACCGGTAGCCTACG	Xhol
(ΔG741 )-orf46.1-	Fwd1	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC	Sall
His	Fwd2	GCGGC <u>GTCGAC</u> GGTGGCGGAGGCACTGGATCCTCAGA	
	Rev	CCCG <u>CTCGAG</u> -CGTATCATATTTCACGTGC	XhoI
(ΔG983)-	Fwd	GCGGC <u>CTCGAG</u> -GGATCCGGAGGGGGTGGTGTCGCC	XhoI
741(MC58) -His			
	Rev	CCCG <u>CTCGAG</u> -TTGCTTGGCGGCAAG	XhoI
(ΔG983)-961c-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGC <u>CTCGAG</u> -GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCG <u>CTCGAG</u> -ACCCAGCTTGTAAGGTTG	XhoI
(ΔG983)-961-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGC <u>CTCGAG</u> -GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCG <u>CTCGAG</u> -CCACTCGTAATTGACGCC	XhoI
(ΔG983)- Orf46.1-	Fwd1	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC	Sall
His	Fwd2	GCGGC <u>GTCGAC</u> GGTGGCGGAGGCACTGGATCCTCAGA	
	Rev	CCCG <u>CTCGAG</u> -CGTATCATATTTCACGTGC	XhoI

<sup>\*</sup> This primer was used as a Reverse primer for all the C terminal fusions of 287 to the His-tag.

§ Forward primers used in combination with the 287-His Reverse primer.

NB – All PCR reactions use strain 2996 unless otherwise specified (e.g. strain MC58)

**[0293]** In all constructs starting with an ATG not followed by a unique *Nhe*l site, the ATG codon is part of the *Nde*l site used for cloning. The constructs made using *Nhe*l as a cloning site at the 5' end (e.g. all those containing 287 at the N-terminus) have two additional codons (GCT AGC) fused to the coding sequence of the antigen.

### Preparation of chromosomal DNA templates

[0294] N.meningitidis strains 2996, MC58, 394.98, 1000 and BZ232 (and others) were grown to exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in 5ml buffer (20% w/v sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1% Na-Sarkosyl, 50µg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two phenol extractions (equilibrated to pH 8) and one CHCl<sub>3</sub>/isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation. The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA concentration was measured by reading OD<sub>260</sub>.

### PCR Amplification

5

10

15

20

25

30

35

40

45

50

[0295] The standard PCR protocol was as follows: 200ng of genomic DNA from 2996, MC581000, or BZ232 strains or 10ng of plasmid DNA preparation of recombinant clones were used as template in the presence of 40μM of each oligonucletide primer, 400-800 μM dNTPs solution, 1x PCR buffer (including 1.5mM MgCl<sub>2</sub>), 2.5 units *Taql* DNA polymerase (using Perkin-Elmer AmpliTaQ, Boerhingher Mannheim Expand<sup>TM</sup> Long Template).

**[0296]** After a preliminary 3 minute incubation of the whole mix at  $95^{\circ}$ C, each sample underwent a two-step amplification: the first 5 cycles were performed using the hybridisation temperature that excluded the restriction enzyme tail of the primer ( $T_{m1}$ ). This was followed by 30 cycles according to the hybridisation temperature calculated for the whole length oligos ( $T_{m2}$ ). Elongation times, performed at  $68^{\circ}$ C or  $72^{\circ}$ C, varied according to the length of the Orf to be amplified. In the case of Orf1 the elongation time, starting from 3 minutes, was increased by 15 seconds each cycle. The cycles were completed with a 10 minute extension step at  $72^{\circ}$ C.

[0297] The amplified DNA was either loaded directly on a 1% agarose gel. The DNA fragment corresponding to the band of correct size was purified from the gel using the Qiagen Gel Extraction Kit, following the manufacturer's protocol.

### Digestion of PCR fragments and of the cloning vectors

[0298] The purified DNA corresponding to the amplified fragment was digested with the appropriate restriction enzymes for cloning into pET-21b+, pET22b+ or pET-24b+. Digested fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted with either H<sub>2</sub>O or 10mM Tris, pH 8.5. Plasmid vectors were digested with the appropriate restriction enzymes, loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit.

## Cloning

10

30

35

40

50

55

**[0299]** The fragments corresponding to each gene, previously digested and purified, were ligated into pET21 b+, pET22b+ or pET-24b+. A molar ratio of 3:1 fragment/vector was used with T4 DNA ligase in the ligation buffer supplied by the manufacturer.

[0300] Recombinant plasmid was transformed into competent *E.coli* DH5 or HB101 by incubating the ligase reaction solution and bacteria for 40 minutes on ice, then at 37°C for 3 minutes.

**[0301]** This was followed by the addition of  $800\mu$ I LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately  $200\mu$ I of the supernatant and plated onto LB ampicillin (100mg/ml) agar.

[0302] Screening for recombinant clones was performed by growing randomly selected colonies overnight at  $37^{\circ}$ C in 4.0ml of LB broth +  $100\mu$ g/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately  $1\mu$ g of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

### Expression

[0303] After cloning each gene into the expression vector, recombinant plasmids were transformed into E.coli strains suitable for expression of the recombinant protein.  $1\mu I$  of each construct was used to transform E.coli BL21-DE3 as described above. Single recombinant colonies were inoculated into 2mILB+Amp ( $100\mu g/mI$ ), incubated at  $37^{\circ}C$  overnight, then diluted 1:30 in 20mI of LB+Amp ( $100\mu g/mI$ ) in 100mI flasks, to give an  $OD_{600}$  between 0.1 and 0.2. The flasks were incubated at  $30^{\circ}C$  or at  $37^{\circ}C$  in a gyratory water bath shaker until  $OD_{600}$  indicated exponential growth suitable for induction of expression (0.4-0.8 OD). Protein expression was induced by addition of 1.0mM IPTG. After 3 hours incubation at  $30^{\circ}C$  or  $37^{\circ}C$  the  $OD_{600}$  was measured and expression examined. 1.0mI of each sample was centrifuged in a microfuge, the pellet resuspended in PBS and analysed by SDS-PAGE and Coomassie Blue staining.

### Gateway cloning and expression

[0304] Sequences labelled GATE were cloned and expressed using the GATEWAY Cloning Technology (GIBCO-BRL). Recombinational cloning (RC) is based on the recombination reactions that mediate the integration and excision of phage into and from the *E.coli* genome, respectively. The integration involves recombination of the *attP* site of the phage DNA within the *attB* site located in the bacterial genome (BP reaction) and generates an integrated phage genome flanked by *attL* and *attR* sites. The excision recombines *attL* and *attR* sites back to *attP* and *attB* sites (LR reaction). The integration reaction requires two enzymes [the phage protein Integrase (Int) and the bacterial protein integration host factor (IHF)] (BP clonase). The excision reaction requires Int, IHF, and an additional phage enzyme, Excisionase (Xis) (LR clonase). Artificial derivatives of the 25-bp bacterial *attB* recombination site, referred to as B 1 and B2, were added to the 5' end of the primers used in PCR reactions to amplify Neisserial ORFs. The resulting products were BP cloned into a "Donor vector" containing complementary derivatives of the phage *attP* recombination site (P1 and P2) using BP clonase. The resulting "Entry clones" contain ORFs flanked by derivatives of the *attL* site (LI and L2) and were subcloned into expression "destination vectors" which contain derivatives of the *attL*-compatible *attR* sites (R1 and R2) using LR clonase. This resulted in "expression clones" in which ORFs are flanked by B1 and B2 and fused in frame to the GST or His N terminal tags.

[0305] The *E. coli* strain used for GATEWAY expression is BL21-SI. Cells of this strain are induced for expression of the T7 RNA polymerase by growth in medium containing salt (0.3 M NaCl).

[0306] Note that this system gives N-terminus His tags.

#### Preparation of membrane proteins.

10

30

35

40

45

50

[0307] Fractions composed principally of either inner, outer or total membrane were isolated in order to obtain recombinant proteins expressed with membrane-localisation leader sequences. The method for preparation of membrane fractions, enriched for recombinant proteins, was adapted from Filip et. al. [J.Bact. (1973) 115:717-722] and Davies et. al. [J.Immunol.Meth. (1990) 143:215-225]. Single colonies harbouring the plasmid of interest were grown overnight at  $37^{\circ}$ C in 20 ml of LB/Amp (100  $\mu$ g/ml) liquid culture. Bacteria were diluted 1:30 in 1.0 L of fresh medium and grown at either 30 °C or  $37^{\circ}$ C until the OD<sub>550</sub> reached 0.6-0.8. Expression of recombinant protein was induced with IPTG at a final concentration of 1.0 mM. After incubation for 3 hours, bacteria were harvested by centrifugation at 8000g for 15 minutes at 4 °C and resuspended in 20 ml of 20 mM Tris-HCl (pH 7.5) and complete protease inhibitors (Boehringer-Mannheim). All subsequent procedures were performed at 4 °C or on ice.

[0308] Cells were disrupted by sonication using a Branson Sonifier 450 and centrifuged at 5000g for 20 min to sediment unbroken cells and inclusion bodies. The supernatant, containing membranes and cellular debris, was centrifuged at 50000g (Beckman Ti50, 29000rpm) for 75 min, washed with 20 mM Bis-tris propane (pH 6.5), 1.0 M NaCl, 10% (v/v) glycerol and sedimented again at 50000g for 75 minutes. The pellet was resuspended in 20mM Tris-HCl (pH 7.5), 2.0% (v/v) Sarkosyl, complete protease inhibitor (1.0 mM EDTA, final concentration) and incubated for 20 minutes to dissolve inner membrane. Cellular debris was pelleted by centrifugation at 5000g for 10 min and the supernatant centrifuged at 75000g for 75 minutes (Beckman Ti50, 33000rpm). Proteins 008L and 519L were found in the supernatant suggesting inner membrane localisation. For these proteins both inner and total membrane fractions (washed with NaCl as above) were used to immunise mice. Outer membrane vesicles obtained from the 75000g pellet were washed with 20 mM Tris-HCl (pH 7.5) and centrifuged at 75000g for 75 minutes or overnight. The OMV was finally resuspended in 500  $\mu$ l of 20 mM Tris-HCl (pH 7.5), 10% v/v glycerol. Orf1L and Orf40L were both localised and enriched in the outer membrane fraction which was used to immunise mice. Protein concentration was estimated by standard Bradford Assay (Bio-Rad), while protein concentration of inner membrane fraction was determined with the DC protein assay (Bio-Rad). Various fractions from the isolation procedure were assayed by SDS-PAGE.

### Purification of His-tagged proteins

[0309] Various forms of 287 were cloned from strains 2996 and MC58. They were constructed with a C-terminus Histagged fusion and included a mature form (aa 18-427), constructs with deletions ( $\Delta$ 1,  $\Delta$ 2,  $\Delta$ 3 and  $\Delta$ 4) and clones composed of either B or C domains. For each clone purified as a His-fusion, a single colony was streaked and grown overnight at 37°C on a LB/Amp (100 μg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 μg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 1.0 L LB/Amp (100 µg/ml) liquid medium and allowed to grow at the optimal temperature (30 or 37°C) until the OD<sub>550</sub> reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000g for 15 min at 4°C. The bacterial pellet was resuspended in 7.5 ml of either (i) cold buffer A (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (10mM Tris-HCl, 100 mM phosphate buffer, pH 8.8 and, optionally, 8M urea) for insoluble proteins. Proteins purified in a soluble form included 287-His,  $\Delta 1$ ,  $\Delta 2$ ,  $\Delta 3$  and  $\Delta 4287$ -His, Δ4287MC58-His, 287c-His and 287cMC58-His. Protein 287bMC58-His was insoluble and purified accordingly. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6 M guanidine hydrochloride, 100 mM phosphate buffer, 10 mM Tris- HCl, pH 7.5 and treated with 10 passes of a Dounce homogenizer. The homogenate was centrifuged at 13000g for 30 min and the supernatant retained. Supernatants for both soluble and insoluble preparations were mixed with 150μl Ni<sup>2+</sup>-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to the manufacturer's protocol. The batch-wise preparation was centrifuged at 700g for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD<sub>280</sub> of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, 20mM imidazole, pH 8.0) or (ii) buffer D (10mM Tris-HCl, 100mM phosphate buffer, pH 6.3 and, optionally, 8M urea) until OD<sub>280</sub> of the flow-through reached 0.02-0.01. The His-fusion protein was eluted by addition of 700μl of either (i) cold elution buffer A (300 mM NaCl, 50mM phosphate buffer, 250 mM imidazole, pH 8.0) or (ii) elution buffer B (10 mM Tris-HCl, 100 mM phosphate buffer, pH 4.5 and, optionally, 8M urea) and fractions collected until the  $OD_{280}$  indicated all the recombinant protein was obtained. 20µl aliquots of each elution fraction were analysed by SDS-PAGE. Protein concentrations were estimated using the Bradford assay.

#### Renaturation of denatured His-fusion proteins.

[0310] Denaturation was required to solubilize 287bMC8, so a renaturation step was employed prior to immunisation. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10% v/v. The proteins were diluted to 200  $\mu$ g/ml using dialysis buffer I (10% v/v glycerol, 0.5M arginine, 50 mM phosphate buffer, 5.0 mM reduced glutathione, 0.5 mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% v/v glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C. Protein concentration was estimated using the formula:

Protein  $(mg/ml) = (1.55 \times OD_{280}) - (0.76 \times OD_{260})$ 

## 15 Amino acid sequence analysis.

**[0311]** Automated sequence analysis of the NH<sub>2</sub>-terminus of proteins was performed on a Beckman sequencer (LF 3000) equipped with an on-line phenylthiohydantoin-amino acid analyser (System Gold) according to the manufacturer's recommendations.

#### **Immunization**

10

25

30

35

40

45

50

[0312] Balb/C mice were immunized with antigens on days 0, 21 and 35 and sera analyzed at day 49.

#### Sera analysis - ELISA

[0313] The acapsulated MenB M7 and the capsulated strains were plated on chocolate agar plates and incubated overnight at 37°C with 5% CO2. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.4-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and bacteria were washed twice with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at 37°C and then overnight at 4°C with stirring. 100μl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS), 200µl of saturation buffer (2.7% polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200µl of diluted sera (Dilution buffer: 1% BSA, 0.1 % Tween-20, 0.1 % NaN<sub>3</sub> in PBS) were added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 100µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100 µl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-phenildiamine and  $10\mu$ l of H<sub>2</sub>O<sub>2</sub>) were added to each well and the plates were left at room temperature for 20 minutes.  $100\mu$ l 12.5%  $H_2SO_4$  was added to each well and  $OD_{490}$  was followed. The ELISA titers were calculated abitrarely as the dilution of sera which gave an OD<sub>490</sub> value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with  $OD_{490}$  of 0.4 was higher than 1:400.

## Sera analysis - FACS Scan bacteria binding assay

[0314] The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C with 5% CO<sub>2</sub>. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN<sub>3</sub>) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD<sub>620</sub> of 0.05.  $100\mu$ l bacterial cells were added to each well of a Costar 96 well plate.  $100\mu$ l of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of  $200\mu$ l/well of blocking buffer in each well.  $100\mu$ l of R-Phicoerytrin conjugated F(ab)<sub>2</sub> goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of  $200\mu$ l/well of blocking buffer. The supernatant was aspirated and cells resuspended in  $200\mu$ l/

well of PBS, 0.25% formaldehyde. Samples were transferred to FACScan tubes and read. The condition for FACScan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

## Sera analysis - bactericidal assay

10

20

30

35

40

45

50

55

[0315] N. meningitidis strain 2996 was grown overnight at  $37^{\circ}$ C on chocolate agar plates (starting from a frozen stock) with 5% CO<sub>2</sub>. Colonies were collected and used to inoculate 7ml Mueller-Hinton broth, containing 0.25% glucose to reach an OD<sub>620</sub> of 0.05-0.08. The culture was incubated for approximately 1.5 hours at 37 degrees with shacking until the OD<sub>620</sub> reached the value of 0.23-0.24. Bacteria were diluted in 50mM Phosphate buffer pH 7.2 containing 10mM MgCl<sub>2</sub>, 10mM CaCl<sub>2</sub> and 0.5% (w/v) BSA (assay buffer) at the working dilution of  $10^5$  CFU/ml. The total volume of the final reaction mixture was 50  $\mu$ l with 25  $\mu$ l of serial two fold dilution of test serum, 12.5  $\mu$ l of bacteria at the working dilution, 12.5  $\mu$ l of baby rabbit complement (final concentration 25%).

[0316] Controls included bacteria incubated with complement serum, immune sera incubated with bacteria and with complement inactivated by heating at  $56^{\circ}$ C for 30'. Immediately after the addition of the baby rabbit complement,  $10\mu$ l of the controls were plated on Mueller-Hinton agar plates using the tilt method (time 0). The 96-wells plate was incubated for 1 hour at 37°C with rotation.  $7\mu$ l of each sample were plated on Mueller-Hinton agar plates as spots, whereas  $10\mu$ l of the controls were plated on Mueller-Hinton agar plates using the tilt method (time 1). Agar plates were incubated for 18 hours at 37 degrees and the colonies corresponding to time 0 and time 1 were counted.

### Sera analysis - western blots

[0317] Purified proteins (500ng/lane), outer membrane vesicles ( $5\mu g$ ) and total cell extracts ( $25\mu g$ ) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3% Tris base, 1.44% glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

[0318] The OMVs were prepared as follows: *N. meningitidis* strain 2996 was grown overnight at 37 degrees with 5% CO<sub>2</sub> on 5 GC plates, harvested with a loop and resuspended in 10 ml of 20mM Tris-HCl pH 7.5, 2 mM EDTA. Heat inactivation was performed at 56°C for 45 minutes and the bacteria disrupted by sonication for 5 minutes on ice (50% duty cycle, 50% output, Branson sonifier 3 mm microtip). Unbroken cells were removed by centrifugation at 5000g for 10 minutes, the supernatant containing the total cell envelope fraction recovered and further centrifuged overnight at 50000g at the temperature of 4°C. The pellet containing the membranes was resuspended in 2% sarkosyl, 20mM Tris-HCl pH 7.5, 2 mM EDTA and incubated at room temperature for 20 minutes to solubilise the inner membranes. The suspension was centrifuged at 10000g for 10 minutes to remove aggregates, the supernatant was further centrifuged at 50000g for 3 hours. The pellet, containing the outer membranes was washed in PBS and resuspended in the same buffer. Protein concentration was measured by the D.C. Bio-Rad Protein assay (Modified Lowry method), using BSA as a standard.

[0319] Total cell extracts were prepared as follows: *N. meningitidis* strain 2996 was grown overnight on a GC plate, harvested with a loop and resuspended in 1 ml of 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30 minutes.

#### 961 domain studies

[0320] Cellular fractions preparation Total lysate, periplasm, supernatant and OMV of E.coli clones expressing different domains of 961 were prepared using bacteria from over-night cultures or after 3 hours induction with IPTG. Briefly, the periplasm were obtained suspending bacteria in saccarose 25% and Tris 50mM (pH 8) with polimixine 100 $\mu$ g/ml. After 1hr at room temperature bacteria were centrifuged at 13000rpm for 15 min and the supernatant were collected. The culture supernatant were filtered with 0.2 $\mu$ m and precipitated with TCA 50% in ice for two hours. After centrifugation (30 min at 13000 rp) pellets were rinsed twice with ethanol 70% and suspended in PBS. The OMV preparation was performed as previously described. Each cellular fraction were analyzed in SDS-PAGE or in Western Blot using the polyclonal anti-serum raised against GST-961.

[0321] Adhesion assay Chang epithelial cells (Wong-Kilbourne derivative, clone 1-5c-4, human conjunctiva) were maintained in DMEM (Gibco) supplemented with 10% heat-inactivated FCS, 15mM L-glutamine and antibiotics.

[0322] For the adherence assay, sub-confluent culture of Chang epithelial cells were rinsed with PBS and treated with

trypsin-EDTA (Gibco), to release them from the plastic support. The cells were then suspended in PBS, counted and dilute in PBS to 5x10<sup>5</sup> cells/ml.

[0323] Bacteria from over-night cultures or after induction with IPTG, were pelleted and washed twice with PBS by centrifuging at 13000 for 5 min. Approximately 2-3x10<sup>8</sup> (cfu) were incubated with 0.5 mg/ml FITC (Sigma) in 1ml buffer containing 50mM NaHCO<sub>3</sub> and 100mM NaCl pH 8, for 30 min at room temperature in the dark. FITC-labeled bacteria were wash 2-3 times and suspended in PBS at 1-1.5x10<sup>9</sup>/ml. 200 $\mu$ l of this suspension (2-3x10<sup>8</sup>) were incubated with 200 $\mu$ l (1x10<sup>5</sup>) epithelial cells for 30min a 37°C. Cells were than centrifuged at 2000rpm for 5 min to remove non-adherent bacteria, suspended in 200 $\mu$ l of PBS, transferred to FACScan tubes and read

Annex to the application documents - subsequently filed sequences listing

[0324]

## SEQUENCE LISTING <110> Chiron SRL Heterologous Expression of Neisserial Proteins <120> P044747EP <130> <140> 06076718.3 <141> 2001-02-28 10 0004695.3 <150> <151> 2000-02-28 0027675.8 <150> <151> 2000-11-13 15 <160> 620 SeqWin99, version 1.02 <170> <210> <211> 441 20 <212> PRT <213> Neisseria meningitidis <400> Met Lys Lys Tyr Leu Phe Arg Ala Ala Leu Tyr Gly Ile Ala Ala 1 1 5 10 15 25 Ile Leu Ala Ala Cys Gln Ser Lys Ser Ile Gln Thr Phe Pro Gln Pro 20 25 30Asp Thr Ser Val Ile Asn Gly Pro Asp Arg Pro Val Gly Ile Pro Asp 40 4530 Pro Ala Gly Thr Thr Val Gly Gly Gly Ala Val Tyr Thr Val Val 50 55 Pro His Leu Ser Leu Pro His Trp Ala Ala Gln Asp Phe Ala Lys Ser 65 70 75 80 Leu Gln Ser Phe Arg Leu Gly Cys Ala Asn Leu Lys Asn Arg Gln Gly 85 90 95 35 Trp Gln Asp Val Cys Ala Gln Ala Phe Gln Thr Pro Val His Ser Phe 100 105 110Gln Ala Lys Gln Phe Phe Glu Arg Tyr Phe Thr Pro Trp Gln Val Ala 115 120 125 40 Gly Asn Gly Ser Leu Ala Gly Thr Val Thr Gly Tyr Tyr Glu Pro Val 130 140 Leu Lys Gly Asp Asp Arg Arg Thr Ala Gln Ala Arg Phe Pro Ile Tyr 145 150 155 45 Gly Ile Pro Asp Asp Phe Ile Ser Val Pro Leu Pro Ala Gly Leu Arg $165 \hspace{1cm} 170 \hspace{1cm} 175$ Ser Gly Lys Ala Leu Val Arg Ile Arg Gln Thr Gly Lys Asn Ser Gly 180 185 50 Thr Ile Asp Asn Thr Gly Gly Thr His Thr Ala Asp Leu Ser Arg Phe 195 200 205 Pro Ile Thr Ala Arg Thr Thr Ala Ile Lys Gly Arg Phe Glu Gly Ser 210 215 220

Arg Phe Leu Pro Tyr His Thr Arg Asn Gln Ile Asn Gly Gly Ala Leu 225 230 240 Asp Gly Lys Ala Pro Ile Leu Gly Tyr Ala Glu Asp Pro Val Glu Leu 245 250 255 Phe Phe Met His Ile Gln Gly Ser Gly Arg Leu Lys Thr Pro Ser Gly 260 265 270Lys Tyr Ile Arg Ile Gly Tyr Ala Asp Lys Asn Glu His Pro Tyr Val 275 280 285 10 Ser Ile Gly Arg Tyr Met Ala Asp Lys Gly Tyr Leu Lys Leu Gly Gln 290 295 300 Thr Ser Met Gln Gly Ile Lys Ala Tyr Met Arg Gln Asn Pro Gln Arg 305 310 315 15 Leu Ala Glu Val Leu Gly Gln Asn Pro Ser Tyr Ile Phe Phe Arg Glu 325 330 335 Leu Ala Gly Ser Ser Asn Asp Gly Pro Val Gly Ala Leu Gly Thr Pro  $340 \hspace{1cm} 345$ 20 Leu Met Gly Glu Tyr Ala Gly Ala Val Asp Arg His Tyr Ile Thr Leu 355 360 365 Gly Ala Pro Leu Phe Val Ala Thr Ala His Pro Val Thr Arg Lys Ala 370 380 25 Leu Asn Arg Leu Ile Met Ala Gln Asp Thr Gly Ser Ala Ile Lys Gly 385 390 395 Ala Val Arg Val Asp Tyr Phe Trp Gly Tyr Gly Asp Glu Ala Gly Glu 405 410 415 30 Leu Ala Gly Lys Gln Lys Thr Thr Gly Tyr Val Trp Gln Leu Leu Pro 420 425 430 Asn Gly Met Lys Pro Glu Tyr Arg Pro 435 440 35 <210> 420 <211> <212> PRT Neisseria meningitidis 40 Gln Ser Lys Ser Ile Gln Thr Phe Pro Gln Pro Asp Thr Ser Val Ile 1 5 10 Asn Gly Pro Asp Arg Pro Val Gly Ile Pro Asp Pro Ala Gly Thr Thr 20 25 30 Val Gly Gly Gly Ala Val Tyr Thr Val Val Pro His Leu Ser Leu 35 40 45 Pro His Trp Ala Ala Gln Asp Phe Ala Lys Ser Leu Gln Ser Phe Arg 50 55 60 Leu Gly Cys Ala Asn Leu Lys Asn Arg Gln Gly Trp Gln Asp Val Cys 65 70 75 80 50 Ala Gln Ala Phe Gln Thr Pro Val His Ser Phe Gln Ala Lys Gln Phe 85 90 95 Phe Glu Arg Tyr Phe Thr Pro Trp Gln Val Ala Gly Asn Gly Ser Leu 100 105 11055

```
Ala Gly Thr Val Thr Gly Tyr Tyr Glu Pro Val Leu Lys Gly Asp Asp 115 120 125
            Arg Arg Thr Ala Gln Ala Arg Phe Pro Ile Tyr Gly Ile Pro Asp Asp 130 140
            Phe Ile Ser Val Pro Leu Pro Ala Gly Leu Arg Ser Gly Lys Ala Leu
145 150 155
            Val Arg Ile Arg Gln Thr Gly Lys Asn Ser Gly Thr Ile Asp Asn Thr 165 170 175
10
            Gly Gly Thr His Thr Ala Asp Leu Ser Arg Phe Pro Ile Thr Ala Arg
180 185 190
            Thr Thr Ala Ile Lys Gly Arg Phe Glu Gly Ser Arg Phe Leu Pro Tyr
195 200 205
15
            His Thr Arg Asn Gln Ile Asn Gly Gly Ala Leu Asp Gly Lys Ala Pro
210 215 220
            Ile Leu Gly Tyr Ala Glu Asp Pro Val Glu Leu Phe Phe Met His Ile
225 230 235 240
20
            Gln Gly Ser Gly Arg Leu Lys Thr Pro Ser Gly Lys Tyr Ile Arg Ile
245 250 255
            Gly Tyr Ala Asp Lys Asn Glu His Pro Tyr Val Ser Ile Gly Arg Tyr
260 265 270
25
            Met Ala Asp Lys Gly Tyr Leu Lys Leu Gly Gln Thr Ser Met Gln Gly 275 280 285
            Ile Lys Ala Tyr Met Arg Gln Asn Pro Gln Arg Leu Ala Glu Val Leu
290 295 300
30
            Gly Gln Asn Pro Ser Tyr Ile Phe Phe Arg Glu Leu Ala Gly Ser Ser 305 310 315
            Asn Asp Gly Pro Val Gly Ala Leu Gly Thr Pro Leu Met Gly Glu Tyr
325 330 335
35
            Ala Gly Ala Val Asp Arg His Tyr Ile Thr Leu Gly Ala Pro Leu Phe 340 350
            Val Ala Thr Ala His Pro Val Thr Arg Lys Ala Leu Asn Arg Leu Ile
355 360 365
            Met Ala Gln Asp Thr Gly Ser Ala Ile Lys Gly Ala Val Arg Val Asp 370 380
40
            Tyr Phe Trp Gly Tyr Gly Asp Glu Ala Gly Glu Leu Ala Gly Lys Gln 385 395 400
            Lys Thr Thr Gly Tyr Val Trp Gln Leu Leu Pro Asn Gly Met Lys Pro
405 410 415
45
            Glu Tyr Arg Pro
420
             <210>
50
             <211>
<212>
                      440
                      PRT
                      Artificial Sequence
             <213>
             <220>
             <223>
                      919
55
             <400>
                      3
```

	Met 1	Lys	Thr	Phe	Phe 5	Lys	Thr	Leu	Ser	А1а 10	Ala	Ala	Leu	Αla	Leu 15	Ile
5	Leu	Ala	Ala	Cys 20	Gln	Ser	Lys	Ser	Ile 25	Gln	Thr	Phe	Pro	G]n 30	Pro	Asp
	Thr	Ser	Va1 35	Ile	Asn	Gly	Pro	Asp 40	Arg	Pro	٧al	Gly	11e 45	Pro	Asp	Pro
10	Ala	G]y 50	Thr	Thr	۷a٦	Gly	G]y 55	Gly	Gly	Ala	val	Tyr 60	Thr	۷al	۷al	Pro
	His 65	Leu	Ser	Leu	Pro	His 70	Trp	Ala	Ala	Gln	Asp 75	Phe	Ala	Lys	Ser	Leu 80
15	Gln	Ser	Phe	Arg	Leu 85	Gly	Cys	Ala	Asn	Leu 90	Lys	Asn	Arg	Gln	Gly 95	Тгр
	Gln	Asp	۷a٦	Cys 100	Ala	Gln	Ala	Phe	Gln 105	Thr	Pro	٧a٦	His	Ser 110	Phe	Gln
20	Ala	Lys	G]n 115	Phe	Phe	Glu	Arg	Tyr 120	Phe	Thr	Pro	Trp	Gln 125	val	Ala	Gly
	Asn	Gly 130	Ser	Leu	Ala	Gly	Thr 135	٧a٦	Thr	Gly	Tyr	Tyr 140	Glu	Pro	val	Leu
25	Lys 145	Gly	Asp	Asp	Arg	Arg 150	Thr	Ala	Gln	Ala	Arg 155	Phe	Pro	Ile	Tyr	Gly 160
	Ile	Pro	Asp	Asp	Phe 165	Ile	Ser	val	Pro	Leu 170	Pro	Ala	Gly	Leu	Arg 175	Ser
30	Gly	Lys	Ala	Leu 180	۷a۱	Arg	Ile	Arg	Gln 185	Thr	Gly	Lys	Asn	Ser 190	Gly	Thr
	Ile	Asp	Asn 195	Thr	Gly	Gly	Thr	His 200	Thr	Ala	Asp.	Leu	Ser 205	Arg	Phe	Pro
0.5	Ile	Thr 210	Ala	Arg	Thr	Thr	Ala 215	Ile	Lys	Gly	Arg	Phe 220	Glu	Gly	Ser	Arg
35	Phe 225	Leu	Pro	Tyr	His	Thr 230	Arg	Asn	Gln	Ile	Asn 235	Gly	Gly	Ala	Leu	Asp 240
	Gly	Lys	Ala	Pro	11e 245	Leu	Gly	Tyr	Ala	Glu 250	Asp	Pro	٧a٦	Glu	Leu 255	Phe
40	Phe	Met	His	11e 260	G∏n	Gly	ser	GТу	Arg 265	Leu	Lys	Thr	Pro	Ser 270	Gly	Lys
	туг	Ile	Arg 275	Ile	Gly	Tyr	Ala	Asp 280	Lys	Asn	Glu	His	Pro 285	Tyr	va1	Ser
45	Ile	Gly 290	Arg	Tyr	Met	Ala	Asp 295	Lys	Gly	Tyr	Leu	Lys 300	Leu	Gly	Gln	Thr
	Ser 305	Met	Gln	Gly	Ile	Lys 310	Ser	Tyr	Met	Arg	G]n 315	Asn	Pro	Gln	Arg	Leu 320
50	Ala	Glu	٧a٦	Leu	G]y 325	Gln	Asn	Pro	Ser	Tyr 330	IJе	Phe	Phe	Arg	Glu 335	Leu
	Ala	Gly	Ser	Ser 340	Asn	Asp	Gly	Pro	Val 345	Gly	Ala	Leu	Gly	Thr 350	Pro	Leu
55	Met	Glу	G]u 355	Tyr	Ala	Gly	Ala	va1 360	Asp	Arg	His	Tyr	11e 365	Thr	Leu	Gly

```
Ala Pro Leu Phe Val Ala Thr Ala His Pro Val Thr Arg Lys Ala Leu 370 380
           Asn Arg Leu Ile Met Ala Gln Asp Thr Gly Ser Ala Ile Lys Gly Ala
385 390 395 400
           Val Arg Val Asp Tyr Phe Trp Gly Tyr Gly Asp Glu Ala Gly Glu Leu 405 410 415
          Ala Gly Lys Gln Lys Thr Thr Gly Tyr Val Trp Gln Leu Leu Pro Asn 420 425 430
10
          Gly Met Lys Pro Glu Tyr Arg Pro
435 440
           <210>
           <211>
           <212>
                    PRT
                    Artificial Sequence
           <213>
           <220>
                    907-2.pep
           <223>
20
           <400>
           Glu Arg Arg Arg Leu Leu Val Asn Ile Gln Tyr Glu Ser Ser Arg Ala
1 5 10 15
           Gly Leu Asp Thr Gln Ile Val Leu Gly Leu Ile Glu Val Glu Ser Ala
20 25 30
           Phe Arg Gln Tyr Ala Ile Ser Gly Val Gly Ala Arg Gly Leu Met Gln 35 40
           Val Met Pro Phe Trp Lys Asn Tyr Ile Gly 50 55
30
           <210>
                    60
           <211>
           <212>
                    Artificial Sequence
           <213>
           <220>
35
                    Escherichia coli
           <223>
           Glu Arg Phe Pro Leu Ala Tyr Asn Asp Leu Phe Lys Arg Tyr Thr Ser
1 5 10 15
40
           Gly Lys Glu Ile Pro Gln Ser Tyr Ala Met Ala Ile Ala Arg Gln Glu
20 25 30
           Ser Ala Trp Asn Pro Lys Val Lys Ser Pro Val Gly Ala Ser Gly Leu
35 40 45
45
           Met Gln Ile Met Pro Gly Thr Ala Thr His Thr Val
50 55 60
           <210>
                    6
120
           <211>
           <212>
50
           <213>
                    Artificial Sequence
           <220>
           <223>
                    922.pep
          Val Ala Gln Lys Tyr Gly Val Pro Ala Glu Leu Ile Val Ala Val Ile
1 5 10
```

```
Gly Ile Glu Thr Asn Tyr Gly Lys Asn Thr Gly Ser Phe Arg Val Ala 20 25 30
          Asp Ala Leu Ala Thr Leu Gly Phe Asp Tyr Pro Arg Arg Ala Gly Phe 35 40 45
          Phe Gln Lys Glu Leu Val Glu Leu Leu Lys Leu Ala Lys Glu Glu Gly 50 55 60
          Gly Asp Val Phe Ala Phe Lys Gly Ser Tyr Ala Gly Ala Met Gly Met 65 70 75 80
10
          Pro Gln Phe Met Pro Ser Ser Tyr Arg Lys Trp Ala Val Asp Tyr Asp
85 90 95
          Gly Asp Gly His Arg Asp Ile Trp Gly Asn Val Gly Asp Val Ala Ala 100 \hspace{1.5cm} 105 \hspace{1.5cm} 110
          Ser Val Ala Asn Tyr Met Lys Gln
115 120
          <210>
                   119
          <211>
20
          <212>
                   PRT
                   Artificial Sequence
          <220>
                   Escherichia coli
          <223>
25
          Ala Trp Gln Val Tyr Gly Val Pro Pro Glu Ile Ile Val Gly Ile Ile
1 5 10 15
          Gly Val Glu Thr Arg Trp Gly Arg Val Met Gly Lys Thr Arg Ile Leu
20 25 30
30
          Asp Ala Leu Ala Thr Leu Ser Phe Asn Tyr Pro Arg Arg Ala Glu Tyr 35 40 45
          Phe Ser Gly Glu Leu Glu Thr Phe Leu Leu Met Ala Arg Asp Glu Gln 50 60
35
          Asp Asp Pro Leu Asn Leu Lys Gly Ser Phe Ala Gly Ala Met Gly Tyr 70 75 80
          Gly Gln Phe Met Pro Ser Ser Tyr Lys Gln Tyr Ala Val Asp Phe Ser
85 90 95
40
          Gly Asp Gly His Ile Asn Leu Trp Asp Pro Val Asp Ala Ile Gly Ser 100 105 110
          Val Ala Asn Tyr Phe Lys Ala
115
45
          <210>
                    <u>1</u>94
          <211>
          <212>
                   PRT
                   Artificial Sequence
          <220>
50
          <223>
                   919.pep
          Ala Leu Asp Gly Lys Ala Pro Ile Leu Gly Tyr Ala Glu Asp Pro Val
1 5 10 15
          Glu Leu Phe Phe Met His Ile Gln Gly Ser Gly Arg Leu Lys Thr Pro
20 25 30
55
```

Ser Gly Lys Tyr Ile Arg Ile Gly Tyr Ala Asp Lys Asn Glu His Pro 35 40 45Tyr Val Ser Ile Gly Arg Tyr Met Ala Asp Lys Gly Tyr Leu Lys Leu 50 60 Gly Gln Thr Ser Met Gln Gly Ile Lys Ser Tyr Met Arg Gln Asn Pro 65 70 75 80 Gln Arg Leu Ala Glu Val Leu Gly Gln Asn Pro Ser Tyr Ile Phe Phe 85 90 95 10 Arg Glu Leu Ala Gly Ser Ser Asn Asp Gly Pro Val Gly Ala Leu Gly 100 105 110 Thr Pro Leu Met Gly Glu Tyr Ala Gly Ala Val Asp Arg His Tyr Ile 115 120 125 15 Thr Leu Gly Ala Pro Leu Phe Val Ala Thr Ala His Pro Val Thr Arg 130 135 140 Lys Ala Leu Asn Arg Leu Ile Met Ala Gln Asp Thr Gly Ser Ala Ile 145 150 155 20 Lys Gly Ala Val Arg Val Asp Tyr Phe Trp Gly Tyr Gly Asp Glu Ala 165 170 175 Gly Glu Leu Ala Gly Lys Gln Lys Thr Thr Gly Tyr Val Trp Gln Leu 180 185 190 25 Leu Pro <210> <211> 196 30 <212> PRT Escherichia coli Ala Leu Ser Asp Lys Tyr Ile Leu Ala Tyr Ser Asn Ser Leu Met Asp 1 10 15 35 Asn Phe Ile Met Asp Val Gln Gly Ser Gly Tyr Ile Asp Phe Gly Asp 20 25 30Gly Ser Pro Leu Asn Phe Phe Ser Tyr Ala Gly Lys Asn Gly His Ala 35 40 45 40 Tyr Arg Ser Ile Gly Lys Val Leu Ile Asp Arg Gly Glu Val Lys Lys 50 60 Glu Asp Met Ser Met Gln Ala Ile Arg His Trp Gly Glu Thr His Ser 70 75 80 45 Glu Ala Glu Val Arg Glu Leu Leu Glu Gln Asn Pro Ser Phe Val Phe 85 90 95 Phe Lys Pro Gln Ser Phe Ala Pro Val Lys Gly Ala Ser Ala Val Pro  $100 \hspace{1cm} 105 \hspace{1cm} 110$ Leu Val Gly Arg Ala Ser Val Ala Ser Asp Arg Ser Ile Ile Pro Pro 115 120 125 50 Gly Thr Thr Leu Leu Ala Glu Val Pro Leu Leu Asp Asn Asn Gly Lys 130 140 Phe Asn Gly Gln Tyr Glu Leu Arg Leu Met Val Ala Leu Asp Val Gly 145 150 155 55

	Gly	Ala	Ile	Lys	Gly 165	Gln	His	Phe	Asp	170	Tyr	Gln	Gly	Ile	Gly 175	Pro	
5	Glu	Ala	Gly	ніs 180	Arg	Ala	Gly	Trp	Tyr 185	Asn	His	Tyr	Gly	Arg 190	val	Trp	
	٧a٦	Leu	Lys 195	Thr													
10	<210 <211 <212 <213	l> ?>	10 28 DNA Art	ifici	ial s	Seque	ence										
15	<220 <223	)> }>	olig	gonu	:Teot	tide											
	<400 cgaa		10 ccc g	gtcgg	gtcti	tt ti	tttta	atg									28
20	<210 <211 <212 <213	l> ?>	11 28 DNA Art	ifici	ial s	Seque	ence										
	<220 <223		olig	gonu	leot	tide											
25	<400 gtg0		11 aaa a	aaaag	gacco	ga c	99991	tct									28
30	<210 <211 <212 <213	l> ?>	12 25 DNA Art	ifici	ial s	Seque	ence										
30	<220 <223		olig	gonu	:Teot	tide											
0.5	<400 aacg		12 cgc o	cggtg	gttti	tg g	gtca										25
35	<210 <211 <212 <213	l> ?>	13 25 DNA Art	ifici	ial s	Seque	ence										
40	<220 <223		olig	gonu	leot	tide											
	<400 tttg		13 caa a	aacao	cgg	cg aq	ggcg										25
45	<210 <211 <212 <213	l> ?>	14 26 DNA Art	ifici	ial s	Seque	ence										
50	<220 <223		olig	gonu	leot	tide											
	<400 tgc		14 gca g	gtcgg	gtcgg	gc a	ctaca	a									26
55	<210 <211 <212 <213	l> ?>	15 26 DNA Art	ifici	ial s	Seque	ence										

	<220> <223>	Oligonucleotide	
5	<400> taatgta	15 gtg ccgaccgact gcgccg	26
10	<210> <211> <212> <213>	16 25 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> tgattga	16 ggt gggtagcgcg ttccg	25
	<210> <211> <212> <213>	17 25 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> ggcggaa	17 cgc gctacccacc tcaat	25
25	<210> <211> <212> <213>	18 34 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> ccggaat	18 tct tatgaaaaaa atcatcttcg ccgc	34
35	<210> <211> <212> <213>	19 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> gcccaage	19 ctt ttattgtttg gctgcctcga tt	32
45	<210> <211> <212> <213>	20 37 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50		20 tct tatgtcgccc gatgttaaat cggcgga	37
	<210> <211> <212> <213>	21 32 DNA Artificial Sequence	
<i>55</i>	<220> <223>	Oligonucleotide	

	<400> gcccaa	21 gctt tcaatcctgc tctttttgc cg	32
5	<210> <211> <212> <213>	22 34 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> ccggaat	22 ttct tatgagccaa gatatggcgg cagt	34
15	<210> <211> <212> <213>	23 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> gcccaag	23 gctt tcaatcctgc tctttttgc cg	32
25	<210> <211> <212> <213>	24 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> ccggaat	24 ttct tatgtccgcc gaatccgcaa atca	34
	<210> <211> <212> <213>	25 32 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> gcccaag	25 gctt tcaatcctgc tctttttgc cg	32
40	<210> <211> <212> <213>	26 36 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> ccggaat	26 ttct tatgggaagg gttgatttgg ctaatg	36
50	<210> <211> <212> <213>	27 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> gcccaag	27 gctt tcaatcctgc tcttttttgc cg	32

5	<210> <211> <212> <213>	28 36 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> ccggaat	28 ttct tatgtcagat ttggcaaacg attctt	36
	<210> <211> <212> <213>	29 35 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> gcccaag	29 gctt ttacgtatca tatttcacgt gcttc	35
20	<210> <211> <212> <213>	30 37 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> ccggaa1	30 ttct tatgtcgccc gatgttaaat cggcgga	37
30	<210> <211> <212> <213>	31 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> gcccaag	31 gctt ttacgtatca tatttcacgt gcttc	35
40	<210> <211> <212> <213>	32 36 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> ccggaat	32 ttct tatgcaaagc aagagcatcc aaacct	36
43	<210> <211> <212> <213>	33 30 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> gcccaag	33 gctt ttacgggcgg tattcgggct	30
55	<210> <211> <212>	34 29 DNA	

	<213>	Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> ccggaat	34 tca tatgaaacac tttccatcc	29
10	<210> <211> <212> <213>	35 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> gcccaage	35 ctt ttaccactcg taattgac	28
20	<210> <211> <212> <213>	36 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> ccggaat	36 tca tatggccaca agcgacgac	29
	<210> <211> <212> <213>	37 28 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> gcccaage	37 ctt ttaccactcg taattgac	28
35	<210> <211> <212> <213>	38 29 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> ccggaat	38 tct tatgaaacac tttccatcc	29
45	<210> <211> <212> <213>	39 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> gcccaage	39 ctt tcaacccacg ttgtaaggtt g	31
55	<210> <211> <212> <213>	40 30 DNA Artificial Sequence	
	<220>		

	<223>	Oligonucleotide	
5	<400> ccggaatt	40 tct tatggccaca aacgacgacg	30
	<210> <211> <212> <213>	41 31 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> gcccaago	41 ctt tcaacccacg ttgtaaggtt g	31
15	<210> <211> <212> <213>	42 34 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> ccggaatt	42 cct tatggccacc tacaaagtgg acga	34
25	<210> <211> <212> <213>	43 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> gcccaago	43 ctt ttattgtttg gctgcctcga tt	32
<i>35</i>	<210> <211> <212> <213>	44 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cgcggato	44 ccg ctagccccga tgttaaatcg gc	32
	<210> <211> <212> <213>	45 31 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cccgctcg	45 gag tcaatcctgc tctttttgc c	31
50	<210> <211> <212> <213>	46 32 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	
	<400>	46	

	cgcggat	ccg ctagccaaga tatggcggca gt	32
5	<210> <211> <212> <213>	47 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> cgcggate	47 ccg ctagcgccga atccgcaaat ca	32
15	<210> <211> <212> <213>	48 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide .	
20	<400> cgcgcta	48 gcg gaagggttga tttggctaat gg	32
	<210> <211> <212> <213>	49 34 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> gggaatte	49 cca tatgggcatt tcccgcaaaa tatc	34
30	<210> <211> <212> <213>	50 32 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cccgctc	50 gag ttacgtatca tatttcacgt gc	32
40	<210> <211> <212> <213>	51 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> gggaatte	51 cca tatgggcatt tcccgcaaaa tatc	34
50	<210> <211> <212> <213>	52 33 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
<i>55</i>	<400> cccgctc	52 gag ttattctatg ccttgtgcgg cat	33
	<210>	53	

	<211> <212> <213>	32 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cgcggat	53 ccc atatggccac aagcgacgac ga	32
10	<210> <211> <212> <213>	54 28 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cccgctc	54 gag ttaccactcg taattgac	28
20	<210> <211> <212> <213>	55 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cgcggat	55 ccc atatggccac aaacgacg	28
30	<210> <211> <212> <213>	56 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cccgctc	56 gag tcatttagca atattatctt tgttc	35
	<210> <211> <212> <213>	57 33 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cgcggat	57 ccc atatgaaagc aaacagtgcc gac	33
45	<210> <211> <212> <213>	58 28 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cccgctc	58 gag ttaccactcg taattgac	28
55	<210> <211> <212> <213>	59 28 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> cgcggat	59 tccc atatggccac aaacgacg	28
10	<210> <211> <212> <213>	60 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cccgcto	60 cgag ttaacccacg ttgtaaggt	29
70	<210> <211> <212> <213>	61 33 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cgcggat	61 cccc atatgatgaa acactttcca tcc	33
25	<210> <211> <212> <213>	62 29 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cccgcto	62 cgag ttaacccacg ttgtaaggt	29
35	<210> <211> <212> <213>	63 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cgcggat	63 cccc atatggccac aaacgacg	28
45	<210> <211> <212> <213>	64 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> cccgcto	64 Egag tcagtctgac actgttttat cc	32
	<210> <211> <212> <213>	65 32 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> cgcggate	65 ccg ctagccccga tgttaaatcg gc	32
5	<210> <211> <212> <213>	66 27 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cccgctcg	66 gag ttacgggcgg tattcgg	27
15	<210> <211> <212> <213>	67 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cgcggate	67 ccg ctagccccga tgttaaatcg gc	32
25	<210> <211> <212> <213>	68 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cccgctcg	68 gag ttacgtatca tatttcacgt gc	32
	<210> <211> <212> <213>	69 32 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cgcggate	69 ccg ctagccccga tgttaaatcg gc	32
40	<210> <211> <212> <213>	70 28 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cccgctcg	70 gag ttaccactcg taattgac	28
50	<210> <211> <212> <213>	71 1457 PRT Neisseria meningitidis	
55	<400> Met Lys 1	71 Thr Thr Asp Lys Arg Thr Thr Glu Thr His Arg Lys Ala Pro 5 10 15	
	Lys Thr	Gly Arg Ile Arg Phe Ser Pro Ala Tyr Leu Ala Ile Cys Leu	

Ser Phe Gly Ile Leu Pro Gln Ala Trp Ala Gly His Thr Tyr Phe Gly 35 40 45Ile Asn Tyr Gln Tyr Tyr Arg Asp Phe Ala Glu Asn Lys Gly Lys Phe 50 55 Ala Val Gly Ala Lys Asp Ile Glu Val Tyr Asn Lys Lys Gly Glu Leu 65 75 80 10 Val Gly Lys Ser Met Thr Lys Ala Pro Met Ile Asp Phe Ser Val Val 85 90 95 Ser Arg Asn Gly Val Ala Ala Leu Val Gly Asp Gln Tyr Ile Val Ser 100 105 110Val Ala His Asn Gly Gly Tyr Asn Asn Val Asp Phe Gly Ala Glu Gly 115 120 125 Arg Asn Pro Asp Gln His Arg Phe Thr Tyr Lys Ile Val Lys Arg Asn 130 135 140 20 Asn Tyr Lys Ala Gly Thr Lys Gly His Pro Tyr Gly Gly Asp Tyr His 145 150 155 Met Pro Arg Leu His Lys Phe Val Thr Asp Ala Glu Pro Val Glu Met 165 170 175 Thr Ser Tyr Met Asp Gly Arg Lys Tyr Ile Asp Gln Asn Asn Tyr Pro  $180 \hspace{1cm} 185 \hspace{1cm} 190$ Asp Arg Val Arg Ile Gly Ala Gly Arg Gln Tyr Trp Arg Ser Asp Glu 195 200 205 Asp Glu Pro Asn Asn Arg Glu Ser Ser Tyr His Ile Ala Ser Ala Tyr 210 215 220 30 Ser Trp Leu Val Gly Gly Asn Thr Phe Ala Gln Asn Gly Ser Gly Gly 225 235 240 Gly Thr Val Asn Leu Gly Ser Glu Lys Ile Lys His Ser Pro Tyr Gly 245 250 255 35 Phe Leu Pro Thr Gly Gly Ser Phe Gly Asp Ser Gly Ser Pro Met Phe 260 265 270Ile Tyr Asp Ala Gln Lys Gln Lys Trp Leu Ile Asn Gly Val Leu Gln 275 280 285 40 Thr Gly Asn Pro Tyr Ile Gly Lys Ser Asn Gly Phe Gln Leu Val Arg 290 295 300 Lys Asp Trp Phe Tyr Asp Glu Ile Phe Ala Gly Asp Thr His Ser Val 305 310 315 45 Phe Tyr Glu Pro Arg Gln Asn Gly Lys Tyr Ser Phe Asn Asp Asp Asn 325 330 Asn Gly Thr Gly Lys Ile Asn Ala Lys His Glu His Asn Ser Leu Pro 340 345 50 Asn Arg Leu Lys Thr Arg Thr Val Gln Leu Phe Asn Val Ser Leu Ser 355 360 365 Glu Thr Ala Arg Glu Pro Val Tyr His Ala Ala Gly Gly Val Asn Ser 370 375 380 55 Tyr Arg Pro Arg Leu Asn Asn Gly Glu Asn Ile Ser Phe Ile Asp Glu

	385					390					395					400
5	Gly	Lys	Gly	Glu	Leu 405	Ile	Leu	Thr	Ser	Asn 410	Ile	Asn	Gln	Gly	Ala 415	GТу
	Gly	Leu	Tyr	Phe 420	Gln	Gly	Asp	Phe	Thr 425	۷a٦	Ser	Pro	Glu	Asn 430	Asn	Glu
10	Thr	Trp	G]n 435	Gly	Ala	GТу	۷al	His 440	Ile	Ser	Glu	Asp	Ser 445	Thr	۷a٦	Thr
	Trp	Lys 450	٧a٦	Asn	Gly	va1	Ala 455	Asn	Asp	Arg	Leu	ser 460	Lys	Ile	Gly	Lys
15	Gly 465	Thr	Leu	His	Val	Gln 470	Ala	Lys	Gly	Glu	Asn 475	Gln	Gly	Ser	Ile	Ser 480
	Val	Gly	Asp	Gly	Thr 485	∨al	Ile	Leu	Asp	Gln 490	Gln	Ala	Asp	Asp	Lys 495	Gly
20	Lys	Lys	Gln	Ala 500	Phe	Ser	Glu	Ile	G]y 505	Leu	val	Ser	Gly	Arg 510	Gly	Thr
20	val	Gln	Leu 515	Asn	Ala	Asp	Asn	G1n 520	Phe	Asn	Pro	Asp	Lys 525	Leu	туг	Phe
	Gly	Phe 530	Arg	Gly	Gly	Arg	Leu 535	Asp	Leu	Asn	Gly	ніs 540	Ser	Leu	Ser	Phe
25	His 545	Arg	Ile	Gln	Asn	Thr 550	Asp	Glu	Gly	Ala	Met 555	Ile	val	Asn	His	Asn 560
	Gln	Asp	Lys	Glu	ser 565	Thr	val	⊤hr	IJe	Thr 570	Gly	Asn	Lys	Asp	11e 575	Ala
30	Thr	Thr	Gly	Asn 580	Asn	Asn	Ser	Leu	Asp 585	Ser	Lys	Lys	Glu	11e 590	Ala	Tyr
	Asn	Gly	Trp 595	Phe	Gly	Glu	Lys	Asp 600	Thr	Thr	Lys	Thr	Asn 605	Gly	Arg	Leu
35	Asn	Leu 610	val	Tyr	Gln	Pro	Ala 615	Ala	Glu	Asp	Arg	Thr 620	Leu	Leu	Leu	Ser
	Gly 625	Gly	Thr	Asn	Leu	Asn 630	Gly	Asn	Ile	Thr	G1n 635	Thr	Asn	Gly	Lys	Leu 640
40	Phe	Phe	Ser	Gly	Arg 645	Pro	Thr	Pro	His	Ala 650	Tyr	Asn	His	Leu	Asn 655	Asp
	His	Trp	Ser	G]n 660	Lys	Glu	Gly	Ile	Pro 665	Arg	Gly	Glu	ıle	val 670	Тгр	Asp
45	Asn	Asp	Trp 675	Ile	Asn	Arg	Thr	Phe 680	Lys	Ala	Glu	Asn	Phe 685	Gln	Ile	Lys
	Gly	Gly 690	Gln	Ala	val	val	Ser 695	Arg	Asn	Val	Ala	Lys 700	val	Lys	Gly	Asp
50	Trp 705	His	Leu	Ser	Asn	ніs 710	Ala	Gln	Ala	٧a٦	Phe 715	Gly	val	Ala	Pro	His 720
	Gln	Ser	His	Thr	11e 725	Cys	Thr	Arg	Ser	Asp 730	Trp	Thr	GТу	Leu	Thr 735	Asn
55	Cys	Val	Glu	Lys 740	Thr	Ile	Thr	Asp	Asp 745	Lys	val	Ile	Ala	Ser 750	Leu	Thr
	Lys	Thr	Asp	Ile	Ser	Gly	Asn	٧a٦	Asp	Leu	Ala	Asp	His	Ala	His	Leu

			755					760					765			
5	Asn	Leu 770	Thr	Glу	Leu	Ala	Thr 775	Leu	Asn	Gly	Asn	Leu 780	Ser	Ala	Asn	Gly
	Asp 785	Thr	Arg	туг	Thr	∨a1 790	Ser	His	Asn	Αla	Thr 795	Gln	Asn	Gly	Asn	Leu 800
10	Ser	Leu	Val	Glу	Asn 805	Ala	Gln	Ala	Thr	Phe 810	Asn	Gln	Ala	Thr	Leu 815	Asn
	Gly	Asn	Thr	Ser 820	Ala	Ser	Gly	Asn	Ala 825	Ser	Phe	Asn	Leu	Ser 830	Asp	His
15	Ala	Val	G1n 835	Asn	Gly	Ser	Leu	Thr 840	Leu	Ser	Gly	Asn	Ala 845	Lys	Ala	Asn
	۷al	<b>Ser</b> <b>8</b> 50	His	Ser	Ala	Leu	Asn 855	Gly	Asn	Val	Ser	Leu 860	Ala	Asp	Lys	Ala
20	va1 865	Phe	His	Phe	Glu	Ser 870	Ser	Arg	Phe	Thr	G1y 875	Gln	Ile	Ser	G∃y	Gly 880
					885			Lys		890					895	
25	-			900	_			Asn	905					910		
20			915	-				Ala 920					925			
00		930					935	Ser	_			940				
30	945					950		val			955					960
			-		965		_	Gln		970					975	
35			-	980	_		·	Lys	985					990		
	-		995					1000	)		_		1005	5		
40		1010	)				101	Glu 5 Gln	-	-		1020	)			
	1025	5				1030	)	Lys			1035	5	-			1040
45		_			1045	5		Ser		1050	)				1055	5
				1060	)			Asn	1069	5		_		1070	)	
50	-	•	1075	5		•	·	1080 Glu	)				1089	5		
		1090	)	_	-		109					1100	)			
55	1109	5			-	111(	)	Asp			1115	5				1120
	Jiu	Ly 3	Lys	~ 9	v a i	9111	A I a	~2P	- 9 3	~2h	• • • • •	n a	Leu	nια	-уэ	5 111

		1125	1130	1135
5	Arg Glu Ala Glu 114		a Thr Thr Ala Phe Pro 1145	Arg Ala Arg 1150
	Arg Ala Arg Arg 1155	g Asp Leu Pro Gli 11	n Leu Gln Pro Gln Pro 60 116	
10	Pro Gln Arg Asj 1170	Leu Ile Ser Ar 1175	g Tyr Ala Asn Ser Gly 1180	'Leu Ser Glu
	Phe Ser Ala Thi 1185	Leu Asn Ser Va 1190	l Phe Ala Val Gln Asp 1195	Glu Leu Asp 1200
15	Arg Val Phe Ala	a Glu Asp Arg Arg 1205	g Asn Ala Val Trp Thr 1210	Ser Gly Ile 1215
	Arg Asp Thr Lys		r Gln Asp Phe Arg Ala 1225	Tyr Arg Gln 1230
20	Gln Thr Asp Let 1235	ı Arg Gln Ile Gl 12	y Met Gln Lys Asn Leu 40 124	
	Arg Val Gly Ile 1250	e Leu Phe Ser Hi 1255	s Asn Arg Thr Glu Asr 1260	Thr Phe Asp
25	1265	1270	g Leu Ala His Gly Ala 1275	1280
23		1285	r Ile Gly Ile Ser Ala 1290	1295
	130	00	p Gly Ile Gly Gly Lys 1305	1310
30	1315	133		25
	1330	1335	e Gly Ala Thr Arg Tyr 1340	
35	1345	1350	n Val Asn Ile Ala Thr 1355	1360
		1365	y Ile Lys Ala Asp Tyr 1370	1375
40	138	30	r Pro Tyr Leu Ser Leu 1385	1390
	1395	14		)5
45	1410	1415	g Ser Ala Glu Trp Gly 1420	
	1425	1430	r Leu His Ala Ala Ala 1435 m Ala Cly Ila Lys Loy	1440
50		1445	r Ala Gly Ile Lys Leu 1450	1455
	Trp			
55	<210> 72 <211> 21 <212> PRT <213> Escher	chia coli		

Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala 1 5 10 Thr Val Ala Gln Ala 20 <210> <211> 73 1439 10 Neisseria meningitidis Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala 1 5 10 Thr Val Ala Gln Ala Ala Ser Ala Gly His Thr Tyr Phe Gly Ile Asn 20 25 30 15 Tyr Gln Tyr Tyr Arg Asp Phe Ala Glu Asn Lys Gly Lys Phe Ala Val 35 45 Gly Ala Lys Asp Ile Glu Val Tyr Asn Lys Lys Gly Glu Leu Val Gly 50 60 20 Lys Ser Met Thr Lys Ala Pro Met Ile Asp Phe Ser Val Val Ser Arg 70 75 80 Asn Gly Val Ala Ala Leu Val Gly Asp Gln Tyr Ile Val Ser Val Ala 85 90 95 25 His Asn Gly Gly Tyr Asn Asn Val Asp Phe Gly Ala Glu Gly Arg Asn  $100 \hspace{1cm} 105 \hspace{1cm} 110$ Pro Asp Gln His Arg Phe Thr Tyr Lys Ile Val Lys Arg Asn Asn Tyr 115 120 125 30 Lys Ala Gly Thr Lys Gly His Pro Tyr Gly Gly Asp Tyr His Met Pro 130 135 140 Arg Leu His Lys Phe Val Thr Asp Ala Glu Pro Val Glu Met Thr Ser 150 155 160 35 Tyr Met Asp Gly Arg Lys Tyr Ile Asp Gln Asn Asn Tyr Pro Asp Arg 165 170 175 Val Arg Ile Gly Ala Gly Arg Gln Tyr Trp Arg Ser Asp Glu Asp Glu 180 185 190 40 Pro Asn Asn Arg Glu Ser Ser Tyr His Ile Ala Ser Ala Tyr Ser Trp 195 200 205 Leu Val Gly Gly Asn Thr Phe Ala Gln Asn Gly Ser Gly Gly Gly Thr 210 215 220 45 Val Asn Leu Gly Ser Glu Lys Ile Lys His Ser Pro Tyr Gly Phe Leu 225 230 240 Pro Thr Gly Gly Ser Phe Gly Asp Ser Gly Ser Pro Met Phe Ile Tyr 245 250 255 50 Asp Ala Gln Lys Gln Lys Trp Leu Ile Asn Gly Val Leu Gln Thr Gly 260 265 270 Asn Pro Tyr Ile Gly Lys Ser Asn Gly Phe Gln Leu Val Arg Lys Asp 275 280 285 55 Trp Phe Tyr Asp Glu Ile Phe Ala Gly Asp Thr His Ser Val Phe Tyr

		290					295					300				
5	G] u 305	Pro	Arg	Gln	Asn	Gly 310	Lys	Tyr	Ser	Phe	Asn 315	Asp	Asp	Asn	Asn	G]y 320
	Thr	Gly	Lys	Ile	Asn 325	Ala	Lys	His	Glu	нis 330	Asn	Ser	Leu	Pro	Asn 335	Arg
10	Leu	Lys	Thr	Arg 340	Thr	val	Gln	Leu	Phe 345	Asn	val	ser	Leu	ser 350	Glu	Thr
10	Ala	Arg	Glu 355	Pro	۷a٦	Tyr	His	Ala 360	Ala	Gly	Gly	Val	Asn 365	Ser	Tyr	Arg
	Pro	Arg 370	Leu	Asn	Asn	Gly	Glu 375	Asn	Ile	Ser	Phe	11e 380	Asp	Glu	Gly	Lys
15	G]y 385	Glu	Leu	Ile	Leu	Thr 390	Ser	Asn	Ile	Asn	G]n 395	Gly	Ala	Gly	Gly	Leu 400
	Tyr	Phe	Gln	Gly	Asp 405	Phe	Thr	val	Ser	Pro 410	Glu	Asn	Asn	Glu	Thr 415	Trp
20	Gln	Gly	Ala	G]y 420	Val	ніѕ	Ile	Ser	Glu 425	Asp	Ser	Thr	۷al	Thr 430	Trp	Lys
	val	Asn	Gly 435	٧a٦	Ala	Asn	Asp	Arg 440	Leu	Ser	Lys	Ile	Gly 445	Lys	Gly	Thr
25	Leu	ніs 450	val	Gln	ΑΊа	Lys	Gly 455	Glu	Asn	Gln	Gly	ser 460	ıle	Ser	val	Gly
	Asp 465	Gly	Thr	Val	Ile	Leu 470	Asp	Gln	Gln	Ala	Asp 475	Asp	Lys	Gly	Lys	Lys 480
30	Gln	Ala	Phe	Ser	Glu 485	Ile	Gly	Leu	val	Ser 490	Gly	Arg	Gly	Thr	Val 495	Gln
	Leu	Asn	Ala	Asp 500	Asn	Gln	Phe	Asn	Pro 505	Asp	Lys	Leu	Tyr	Phe 510	Gly	Phe
35	Arg	Gly	Gly 515	Arg	Leu	Asp	Leu	Asn 520	Gly	ніѕ	Ser	Leu	Ser 525	Phe	His	Arg
	Ile	G]n 530	Asn	Thr	Asp	Glu	G]y 535	Ala	Met	Ile	val	Asn 540	His	Asn	Gln	Asp
40	Lys 545	Glu	Ser	Thr	val	Thr 550	Ile	Thr	Gly	Asn	Lys 555	Asp	Ile	Ala	Thr	Thr 560
	Gly	Asn	Asn	Asn	Ser 5 <b>6</b> 5	Leu	Asp	Ser	Lys	Lys 570	Glu	Ile	Ala	Tyr	Asn 575	Gly
45	тгр	Phe	Gly	G1u 580	Lys	Asp	Thr	Thr	Lys 585	Thr	Asn	Gly	Arg	Leu 590	Asn	Leu
	val	Tyr	Gln 595	Pro	Ala	Ala	Glu	Asp 600	Arg	Thr	Leu	Leu	Leu 605	Ser	Gly	Gly
50	Thr	Asn 610	Leu	Asn	G⅂y	Asn	Ile 615	Thr	Gln	Thr	Asn	Gly 620	Lys	Leu	Phe	Phe
	Ser 625	Gly	Arg	Pro	Thr	Pro 630	His	Ala	Tyr	Asn	нis 635	Leu	Asn	Asp	His	Trp 640
55	Ser	Gln	Lys	Glu	G]y 645	ıJe	Pro	Arg	Gly	G1u 650	ıle	٧a٦	тгр	Asp	Asn 655	Asp
	Trp	Ile	Asn	Arg	Thr	Phe	Lys	Аlа	Glu	Asn	Phe	Gln	Пe	Lys	Gly	GТу

				660					665					670		
5	Gln	Ala	Val 675	Val	Ser	Arg	Asn	va1 680	Ala	Lys	٧a٦	Lys	Gly 685	Asp	Тгр	His
	Leu	Ser 690	Asn	His	Ala	Gln	Ala 695	val	Phe	Gly	٧a٦	Ala 700	Pro	His	Gln	Ser
10	His 705	Thr	Ile	Cys	Thr	Arg 710	Ser	Asp	Trp	Thr	Gly 715	Leu	Thr	Asn	Cys	∨a1 720
	Glu	Lys	Thr	Ile	Thr 725	Asp	Asp	Lys	Val	11e 730	Ala	Ser	Leu	Thr	Lys 735	Thr
15	Asp	Ile	Ser	Gly 740	Asn	va1	Asp	Leu	Ala 745	Asp	ніѕ	Ala	His	Leu 750	Asn	Leu
15	Thr	Gly	Leu 755	Ala	Thr	Leu	Asn	Gly 760	Asn	Leu	Ser	Ala	Asn 765	Gly	Asp	Thr
	Arg	Tyr 770	Thr	Val	Ser	His	Asn 775	Ala	Thr	Gln	Asn	Gly 780	Asn	Leu	Ser	Leu
20	∨a1 785	Gly	Asn	Аlа	Gln	А1а 790	Thr	Phe	Asn	Gln	Ala 795	Thr	Leu	Asn	Gly	Asn 800
	Thr	Ser	Ala	Ser	Gly 805	Asn	Ala	Ser	Phe	Asn 810	Leu	Ser	Asp	His	Ala 815	∨al
25	Gln	Asn	Gly	Ser 820	Leu	Thr	Leu	Ser	Gly 825	Asn	Ala	Lys	Ala	Asn 830	٧a٦	ser
	His	Ser	Ala 835	Leu	Asn	Gly	Asn	va1 840	Ser	Leu	Ala	Asp	Lys 845	Ala	Val	Phe
30	His	Phe 850	Glu	Ser	Ser	Arg	Phe 855	Thr	Gly	Gln	Ile	Ser 860	Gly	Gly	Lys	Asp
	Thr 865	Ala	Leu	His	Leu	Lys 870	Asp	Ser	Glu	Trp	Thr 875	Leu	Pro	Ser	Gly	Thr 880
35	Glu	Leu	Gly	Asn	Leu 885	Asn	Leu	Asp	Asn	Ala 890	Thr	Ile	Thr	Leu	Asn 895	Ser
	Ala	Tyr	Arg	ніs 900	Asp	Ala	Ala	Gly	Ala 905	Gln	Thr	Gly	ser	Ala 910	Thr	Asp
40	Ala	Pro	Arg 915	Arg	Arg	Ser	Arg	Arg 920	Ser	Arg	Arg	Ser	Leu 925	Leu	Ser	٧a٦
	Thr	Pro 930	Pro	Thr	ser	va1	G]u 935	Ser	Arg	Phe	Asn	Thr 940	Leu	Thr	val	Asn
45	Gly 945	Lys	Leu	Asn	Gly	G]n 950	Gly	Thr	Phe	Arg	Phe 955	Met	Ser	Glu	Leu	Phe 960
	Gly	Tyr	Arg	Ser	Asp 965	Lys	Leu	Lys	Leu	Ala 970	Glu	Ser	Ser	Glu	G]y 975	Thr
50	Tyr	Thr	Leu	Ala 980	۷a٦	Asn	Asn	Thr	Gly 985	Asn	Glu	Pro	Ala	Ser 990	Leu	Glu
	Gln	Leu	Thr 995	۷a٦	۷a٦	Glu	Glу	Lys 1000		Asn	Lys	Pro	Leu 1005		Glu	Asn
55	Leu	Asn 1010		Thr	Leu	Gln	Asn 101		His	val	Asp	Ala 1020		Ala	Trp	Arg
	Tyr	Gln	Leu	Ile	Ara	Lys	Asp	Gly	Glu	Phe	Arg	Leu	His	Asn	Pro	val

	1025	5				1030	)				1035	5				1040
5	Lys	Glu	Gln	Glu	Leu 1045		Asp	Lys	Leu	Gly 1050		Ala	Glu	Ala	Lys 1059	
	Gln	Ala	Glu	Lys 1060		Asn	Ala	Gln	Ser 1065		Asp	Ala	Leu	lle 1070		Ala
10	Gly	Arg	Asp 1075		٧a٦	Glu	Lys	Thr 1080		Ser	۷al	Ala	Glu 1089		Ala	Arg
	Gln	Ala 1090		Gly	Glu	Asn	val 1095		Ile	Met	Gln	Ala 1100		Glu	Glu	Lys
15	Lys 110		val	Gln	Ala	Asp 111(	Lys )	Asp	Thr	Ala	Leu 1115		Lys	Gln	Arg	Glu 1120
	Ala	Glu	Thr	Arg	Pro 1125		Thr	Thr	Ala	Phe 1130		Arg	Ala	Arg	Arg 1135	
20	Arg	Arg	Asp	Leu 1140		Gln	Leu	Gln	Pro 1145		Pro	Gln	Pro	Gln 115(		Gln
20	Arg	Asp	Leu 115		Ser	Arg	Tyr	Ala 1160		Ser	Gly	Leu	Ser 116		Phe	Ser
	Ala	Thr 1170	Leu )	Asn	Ser	Val	Phe 1175	Ala	۷a٦	Gln	Asp	Glu 1180		Asp	Arg	۷al
25	Phe 1185		Glu	Asp	Arg	Arg 1190	Asn )	Ala	۷a٦	⊤rp	Thr 1199		Gly	Ile	Arg	Asp 1200
	Thr	Lys	His	Tyr	Arg 1205		Gln	Asp	Phe	Arg 121(		Tyr	Arg	Gln	Gln 1215	
30	Asp	Leu	Arg	G]n 1220		Gly	Met	Gln	Lys 1225		Leu	Gly	ser	Gly 1230		val
	Gly	Ile	Leu 1235		ser	His	Asn	Arg 1240		Glu	Asn	Thr	Phe 1245		Asp	Gly
35	Ile	Gly 1250		Ser	Ala	Arg	Leu 1255		His	Gly	Ala	val 1260		Gly	Gln	Tyr
	Gly 1265		Asp	Arg	Phe	Tyr 127(	Ile )	Gly	Ile	Ser	Ala 1275		Ala	Gly	Phe	Ser 1280
40	Ser	Gly	Ser	Leu	Ser 1289		Gly	Ile	Glу	Gly 1290		Ile	Arg	Arg	Arg 1299	
	Leu	His	Tyr	Gly 1300		Gln	Ala	Arg	Tyr 1305		Ala	Gly	Phe	Gly 1310		Phe
45	Gly	Ile	Glu 1319		His	Ile	Gly	Ala 1320		Arg	Tyr	Phe	val 1325	Gln	Lys	Ala
	Asp	Tyr 133(		Tyr	Glu	Asn	Val 1335		Ile	Ala	Thr	Pro 1340		Leu	Ala	Phe
50	Asn 1345		Tyr	Arg	Ala	Gly 1350	Ile )	Lys	Ala	Asp	Tyr 1355		Phe	Lys	Pro	Ala 1360
	Gln	His	Ile	Ser	11e 1365		Pro	Tyr	Leu	Ser 1370		Ser	Tyr	Thr	Asp 1375	
55	Ala	Ser	Gly	Lys 1380	val )	Arg	Thr	Arg	Val 1385		Thr	Ala	val	Leu 1390		Gln
	Asp	Phe	Gly	Lys	Thr	Arg	Ser	Ala	Glu	Trp	Gly	val	Asn	Ala	Glu	Ile

		1395		140	0		1405	
5	Lys Gly 141		Leu Ser	Leu His 1415	Ala Ala	Ala Ala 1420		Pro Gln
	Leu Glu 1425	Ala Gln	His Ser 143		Ile Lys	Leu Gly 1435	Tyr Arg	Trp
10	<210> <211> <212> <213>	74 164 PRT Neisser	ia menin	gitidis				
45	<400> Met Lys 1	74 Lys Asn	Ile Leu 5	Glu Phe	Trp Val	Gly Leu	Phe Val	Leu Ile 15
15	Gly Ala	Ala Ala 20	Val Ala	Phe Leu	Ala Phe 25	Arg Val	Ala Gly 30	Gly Ala
	Ala Phe	Gly Gly 35	Ser Asp	Lys Thr 40	Tyr Ala	Val Tyr	Ala Asp 45	Phe Gly
20	Asp Ile 50	Gly Gly	Leu Lys	Val Asn 55	Ala Pro	Val Lys 60	Ser Ala	Gly Val
	Leu Val 65	Gly Arg	Val Gly 70	Ala Ile	Gly Leu	Asp Pro 75	Lys Ser	Tyr Gln 80
25	Ala Arg	Val Arg	Leu Asp 85	Leu Asp	Gly Lys 90	Tyr Gln	Phe Ser	Ser Asp 95
	Val Ser	Ala Gln 100	Ile Leu	Thr Ser	Gly Leu 105	Leu Gly	Glu Gln 110	Tyr Ile
30	Gly Leu	Gln Gln 115	Gly Gly	Asp Thr 120	Glu Asn	Leu Ala	Ala Gly 125	Asp Thr
	Ile Ser 130		Ser Ser	Ala Met 135	Val Leu	Glu Asn 140	Leu Ile	Gly Lys
35	Phe Met 145	Thr Ser	Phe Ala 150		Asn Ala	Asp Gly 155	Gly Asn	Ala Glu 160
	Lys Ala	Ala Glu						
40	<210> <211> <212> <213>	75 21 PRT Erwinia	carotov	ora				
45	<400> Met Lys 1	75 Tyr Leu	Leu Pro 5	Thr Ala	Ala Ala 10	Gly Leu	Leu Leu	Ala Ala 15
	Gln Pro	Ala Met 20	Ala					
50	<210> <211> <212> <213>	76 608 PRT Neisser	ia menin	gitidis (	ORF46			
55	<400> Leu Gly 1	76 Ile Ser	Arg Lys 5	Ile Ser	Leu Ile 10	Leu Ser	Ile Leu	Ala Val 15
-	Cys Leu	Pro Met	His Ala	His Ala	Ser Asp	Leu Ala	Asn Asp	Ser Phe

25 20 30 Ile Arg Gln Val Leu Asp Arg Gln His Phe Glu Pro Asp Gly Lys Tyr 35 40His Leu Phe Gly Ser Arg Gly Glu Leu Ala Glu Arg Ser Gly His Ile 50 55 Gly Leu Gly Lys Ile Gln Ser His Gln Leu Gly Asn Leu Met Ile Gln 65 75 8010 Gln Ala Ala Ile Lys Gly Asn Ile Gly Tyr Ile Val Arg Phe Ser Asp 85 90 95 His Gly His Glu Val His Ser Pro Phe Asp Asn His Ala Ser His Ser 100 105 11015 Asp Ser Asp Glu Ala Gly Ser Pro Val Asp Gly Phe Ser Leu Tyr Arg 115 120 125 Ile His Trp Asp Gly Tyr Glu His His Pro Ala Asp Gly Tyr Asp Gly
130 140 20 Pro Gln Gly Gly Tyr Pro Ala Pro Lys Gly Ala Arg Asp Ile Tyr 145 150 155 Ser Tyr Asp Ile Lys Gly Val Ala Gln Asn Ile Arg Leu Asn Leu Thr 165 170 25 Asp Asn Arg Ser Thr Gly Gln Arg Leu Ala Asp Arg Phe His Asn Ala  $180 \hspace{1cm} 185 \hspace{1cm} 190$ Gly Ser Met Leu Thr Gln Gly Val Gly Asp Gly Phe Lys Arg Ala Thr  $195 \hspace{1.5cm} 200 \hspace{1.5cm} 205$ Arg Tyr Ser Pro Glu Leu Asp Arg Ser Gly Asn Ala Ala Glu Ala Phe 210 220 30 Asn Gly Thr Ala Asp Ile Val Lys Asn Ile Ile Gly Ala Ala Gly Glu 225 230 235 240 Ile Val Gly Ala Gly Asp Ala Val Gln Gly Ile Ser Glu Gly Ser Asn 245 250 255 35 Ile Ala Val Met His Gly Leu Gly Leu Leu Ser Thr Glu Asn Lys Met  $260 \hspace{1.5cm} 265 \hspace{1.5cm} 270 \hspace{1.5cm}$ Ala Arg Ile Asn Asp Leu Ala Asp Met Ala Gln Leu Lys Asp Tyr Ala 275 280 285 40 Ala Ala Ile Arg Asp Trp Ala Val Gln Asn Pro Asn Ala Ala Gln 290 295 300 Gly Ile Glu Ala Val Ser Asn Ile Phe Met Ala Ala Ile Pro Ile Lys  $305 \hspace{1.5cm} 310 \hspace{1.5cm} 315$ 45 Gly Ile Gly Ala Val Arg Gly Lys Tyr Gly Leu Gly Gly Ile Thr Ala 325 330 335 His Pro Ile Lys Arg Ser Gln Met Gly Ala Ile Ala Leu Pro Lys Gly 340 345 350 50 Lys Ser Ala Val Ser Asp Asn Phe Ala Asp Ala Ala Tyr Ala Lys Tyr 355 360 365 Pro Ser Pro Tyr His Ser Arg Asn Ile Arg Ser Asn Leu Glu Gln Arg 370 380 55 Tyr Gly Lys Glu Asn Ile Thr Ser Ser Thr Val Pro Pro Ser Asn Gly

	385					390					395					400
5	Lys	Asn	<b>v</b> al	Lys	Leu 405	Ala	Asp	Gln	Arg	ніs 410	Pro	Lys	Thr	Gly	Val 415	Pro
	Phe	Asp	Gly	Lys 420	GТу	Phe	Pro	Asn	Phe 425	Glu	Lys	ніѕ	val	Lys 430	Tyr	Asp
10	Thr	Lys	Leu 435	Asp	Ile	Gln	Glu	Leu 440	Ser	Gly	Gly	Gly	11e 445	Pro	Lys	Ala
	Lys	Pro 450	val	Ser	Asp	ΑΊа	Lys 455	Pro	Arg	Тгр	Glu	∨a1 460	Asp	Arg	Lys	Leu
15	Asn 465	Lys	Leu	Thr	Thr	Arg 470	Glu	Gln	۷al	Glu	Lys 475	Asn	Val	Gln	Glu	Ile 480
	Arg	Asn	Gly	Asn	Lys 485	Asn	ser	Asn	Phe	ser 490	Gln	His	Ala	Gln	Leu 495	Glu
20	Arg	Glu	Ile	Asn 500	Lys	Leu	Lys	Ser	Ala 505	Asp	Glu	Ile	Asn	Phe 510	Ala	Asp
	Gly	Met	Gly 515	Lys	Phe	Thr	Asp	Ser 520	Met	Asn	Asp	Lys	Ala 525	Phe	ser	Arg
25	Leu	va1 530	Lys	Ser	۷a٦	Lys	Glu 535	Asn	Gly	Phe	Thr	Asn 540	Pro	val	∨al	Glu
	Tyr 545	Val	Glu	Ile	Asn	Gly 550	Lys	Ala	Tyr	Ile	Val 555	Arg	Gly	Asn	Asn	Arg 560
	۷a٦	Phe	Ala	Ala	Glu 565	Tyr	Leu	Gly	Arg	11e 570	His	Glu	Leu	Lys	Phe 575	Lys
30	Lys	٧a٦	Asp	Phe 580	Pro	٧a٦	Pro	Asn	Thr 585	Ser	Trp	Lys	Asn	Pro 590	Thr	Asp
	۷a٦	Leu	Asn 595	Glu	Ser	Gly	Asn	va1 600	Lys	Arg	Pro	Arg	Туг 605	Arg	Ser	Lys
35	<210 <211 <211 <211	1> 2>	77 584 PRT Art	ifici	ial S	Seque	ence									
40	<220 <223		ORF4	16-2												
	<400 ser 1		77 Leu	Ala	Asn 5	Asp	ser	Phe	Ile	Arg 10	Gln	val	Leu	Asp	Arg 15	Gln
45	His	Phe	Glu	Pro 20	Asp	Gly	Lys	Tyr	His 25	Leu	Phe	Gly	Ser	Arg 30	Gly	Glu
	Leu	Ala	Glu 35	Arg	Ser	Gly	His	Ile 40	Glу	Leu	Glу	Lys	17e 45	Gln	Ser	His
50	Gln	Leu 50	Gly	Asn	Leu	Met	Ile 55	Gln	Gln	Ala	Ala	11e 60	Lys	Gly	Asn	Ile
	Gly 65	Tyr	Ile	Val	Arg	Phe 70	ser	Asp	His	Gly	His 75	Glu	val	His	Ser	Pro 80
55	Phe	Asp	Asn	His	А1а 85	ser	ніѕ	Ser	Asp	Ser 90	Asp	Glu	Аlа	Gly	Ser 95	Pro

Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu His 100 105 110His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Tyr Pro Ala 115 120 125 Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val Ala 130 135 140 Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln Arg 145 150 155 160 10 Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly Val 165 170 175 Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp Arg 180 185 19015 Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val Lys 195 200 205 Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp Ala Val 210 220 20 Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu Gly 225 235 240 Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala Asp 245 250 255 25 Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ala Ile Arg Asp Trp Ala 260 265 270 Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn Ile 275 280 285 30 Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly Lys 290 295 300 Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln Met 305 310 315 32035 Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn Phe  $325 \hspace{1.5cm} 330 \hspace{1.5cm} 335$ Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg Asn 340 345 350 40 Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr Ser 355 360Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp Gln 370 375 380 45 Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro Asn 385 390 400 Phe Glu Lys His Val Lys Tyr Asp Thr Lys Leu Asp Ile Gln Glu Leu 405 410 415 Ser Gly Gly Ile Pro Lys Ala Lys Pro Val Ser Asp Ala Lys Pro 420 425 430 50 Arg Trp Glu Val Asp Arg Lys Leu Asn Lys Leu Thr Thr Arg Glu Gln 435 440 445 Val Glu Lys Asn Val Gln Glu Ile Arg Asn Gly Asn Lys Asn Ser Asn 450 455 460 55

Phe Ser Gln His Ala Gln Leu Glu Arg Glu Ile Asn Lys Leu Lys Ser 465 470 475 480 Ala Asp Glu Ile Asn Phe Ala Asp Gly Met Gly Lys Phe Thr Asp Ser 485 490 495 Met Asn Asp Lys Ala Phe Ser Arg Leu Val Lys Ser Val Lys Glu Asn 500 505 510 Gly Phe Thr Asn Pro Val Val Glu Tyr Val Glu Ile Asn Gly Lys Ala 515 520 525 10 Tyr Ile Val Arg Gly Asn Asn Arg Val Phe Ala Ala Glu Tyr Leu Gly 530 540 Arg Ile His Glu Leu Lys Phe Lys Lys Val Asp Phe Pro Val Pro Asn 545 550 555 15 Thr Ser Trp Lys Asn Pro Thr Asp Val Leu Asn Glu Ser Gly Asn Val 565 570 575 Lys Arg Pro Arg Tyr Arg Ser Lys 580 20 <211> 364 <212> PRT <213> Neisseria meningitidis 25 Met Ser Met Lys His Phe Pro Ala Lys Val Leu Thr Thr Ala Ile Leu 1 5 10 15 Ala Thr Phe Cys Ser Gly Ala Leu Ala Ala Thr Ser Asp Asp Val 20 25 30 30 Lys Lys Ala Ala Thr Val Ala Ile Val Ala Ala Tyr Asn Asn Gly Gln 35 40 45 Glu Ile Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile Gly Glu
50 55 35 Asp Gly Thr Ile Thr Gln Lys Asp Ala Thr Ala Ala Asp Val Glu Ala 65 70 75 80 Asp Asp Phe Lys Gly Leu Gly Leu Lys Lys Val Val Thr Asn Leu Thr 85 90 95 40 Lys Thr Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val Lys Ala 100 105 110Ala Glu Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp Thr Asp 115 120 125 Ala Ala Leu Ala Asp Thr Asp Ala Ala Leu Asp Glu Thr Thr Asn Ala 130 140 45 Leu Asn Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu Thr Lys 145 155 160 Thr Asn Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala Asp Thr 165 170 175 50 Val Asp Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp Ser Leu Asp 180 185 Glu Thr Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala Asn Glu Ala 195 200 205 55

Lys Gln Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala Lys Val Lys 210 215 220 Ala Ala Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Gly Thr Ala 225 230 235 240 Asn Thr Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys Val Thr Asp 245 250 255 Ile Lys Ala Asp Ile Ala Thr Asn Lys Ala Asp Ile Ala Lys Asn Ser 260 265 270 10 Ala Arg Ile Asp Ser Leu Asp Lys Asn Val Ala Asn Leu Arg Lys Glu 275 280 285 Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu Phe Gln 290 295 300 15 Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val Gly Gly Tyr 305 310 315 Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg Phe Thr Glu 325 330 335 20 Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser Ser Gly Ser 340 345 Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp 355 360 25 79 427 <210> <211> <212> PRT Neisseria meningitidis 30 Met Phe Glu Arg Ser Val Ile Ala Met Ala Cys Ile Phe Ala Leu Ser 1 10 15 Ala Cys Gly Gly Gly Gly Gly Ser Pro Asp Val Lys Ser Ala Asp 20 25 30 35 Thr Leu Ser Lys Pro Ala Ala Pro Val Val Ala Glu Lys Glu Thr Glu 35 40 45 Val Lys Glu Asp Ala Pro Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro 50 60 40 Ser Thr Gln Gly Ser Gln Asp Met Ala Ala Val Ser Ala Glu Asn Thr 65 70 80 Gly Asn Gly Gly Ala Ala Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu 85 90 95 45 Gly Pro Gln Asn Asp Met Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln 100 105Thr Gly Asn Asn Gln Pro Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser 115 120 125 Asn Pro Ala Pro Ala Asn Gly Gly Ser Asn Phe Gly Arg Val Asp Leu 130 140 50 Ala Asn Gly Val Leu Ile Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr 145 150 160 His Cys Lys Gly Asp Ser Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu 165 170 175 55

```
Ala Pro Ser Lys Ser Glu Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile 180 \hspace{1cm} 185 \hspace{1cm} 190
            Glu Lys Tyr Lys Lys Asp Gly Lys Ser Asp Lys Phe Thr Asn Leu Val
195 200 205
            Ala Thr Ala Val Gln Ala Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr 210 220
            Lys Asp Lys Ser Ala Ser Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala
225 230 235 240
10
            Arg Ser Arg Arg Ser Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn
245 250 255
            Gln Ala Asp Thr Leu Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly 260 265 270
15
            His Ser Gly Asn Ile Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr 275 280 285
            Tyr Gly Ala Glu Lys Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln
290 295 300
20
            Gly Glu Pro Ala Lys Gly Glu Met Leu Ala Gly Thr Ala Val Tyr Asn
305 310 315
            Gly Glu Val Leu His Phe His Thr Glu Asn Gly Arg Pro Tyr Pro Thr
325 330 335
25
            Arg Gly Arg Phe Ala Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp 340 345
            30
            Lys Ala Ala Ile Asp Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn 370 380
            Gly Gly Gly Asp Val Ser Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu
385 390 395 400
35
            Val Ala Gly Lys Tyr Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly 405 410 415
            Phe Gly Val Phe Ala Gly Lys Lys Glu Gln Asp
420 425
40
            <210>
                     80
                     410
            <211>
            <212>
            <213>
                     Artificial Sequence
            <220>
45
            <223>
                     287untagged
            Cys Gly Gly Gly Gly Gly Ser Pro Asp Val Lys Ser Ala Asp Thr 1 5 10
50
            Leu Ser Lys Pro Ala Ala Pro Val Val Ala Glu Lys Glu Thr Glu Val
20 25 30
            Lys Glu Asp Ala Pro Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser
40
45
            Thr Gln Gly Ser Gln Asp Met Ala Ala Val Ser Ala Glu Asn Thr Gly 50 55
55
```

	Asn 65	Gly	Gly	Аlа	Аlа	Thr 70	Thr	Asp	Lys	Pro	Lys 75	Asn	Glu	Asp	Glu	G]y 80
5	Pro	Gln	Asn	Asp	Met 85	Pro	Gln	Asn	Ser	Аlа 90	Glu	ser	Ala	Asn	G]n 95	Thr
	Gly	Asn	Asn	G]n 100	Pro	Ala	Asp	Ser	Ser 105	Asp	Ser	Ala	Pro	Ala 110	Ser	Asn
10	Pro	Аlа	Pro 115	Ala	Asn	Gly	Gly	Ser 120	Asn	Phe	Gly	Arg	Val 125	Asp	Leu	Ala
	Asn	Gly 130	val	Leu	Ile	Asp	Gly 135	Pro	Ser	Gln	Asn	Ile 140	Thr	Leu	Thr	ніѕ
15	Cys 145	Lys	Gly	Asp	Ser	Cys 150	Asn	Gly	Asp	Asn	Leu 155	Leu	Asp	Glu	Glu	Ala 160
	Pro	Ser	Lys	Ser	Glu 165	Phe	Glu	Asn	Leu	Asn 170	Glu	Ser	Glu	Arg	Ile 175	Glu
20	Lys	Tyr	Lys	Lys 180	Asp	Gly	Lys	Ser	Asp 185	Lys	Phe	Thr	Asn	Leu 190	Val	Ala
	Thr	Ala	∨a1 195	Gln	Аlа	Asn	GТу	Thr 200	Asn	Lys	Tyr	val	11e 205	ıle	туг	Lys
25	Asp	Lys 210	Ser	Ala	Ser	ser	Ser 215	ser	Аlа	Arg	Phe	Arg 220	Arg	Ser	Аlа	Arg
23	Ser 225	Arg	Arg	Ser	Leu	Pro 230	Ala	Glu	Met	Pro	Leu 235	Ile	Pro	Val	As <b>n</b>	Gln 240
	Ala	Asp	Thr	Leu	Ile 245	۷a٦	Asp	Gly	Glu	Ala 250	٧a٦	Ser	Leu	Thr	G]y 255	His
30	Ser	Gly	Asn	11e 260	Phe	Ala	Pro	Glu	G]y 265	Asn	Tyr	Arg	Tyr	Leu 270	Thr	туг
	Gly	Аlа	Glu 275	Lys	Leu	Pro	Gly	G1y 280	ser	Tyr	Ala	Leu	Arg 285	val	Gln	Gly
35	Glu	Pro 290	Ala	Lys	Gly	Glu	Met 295	Leu	Ala	Gly	Thr	Ala 300	val	Tyr	Asn	Gly
	G1u 305	۷al	Leu	ніѕ	Phe	His 310	Thr	Glu	Asn	Gly	Arg 315	Pro	Tyr	Pro	Thr	Arg 320
40	Gly	Arg	Phe	Ala	Ala 325	Lys	۷al	Asp	Phe	G]y 330	Ser	Lys	Ser	۷al	Asp 335	Gly
	Ile	Ile	Asp	Ser 340	Gly	Asp	Asp	Leu	ніs 345	Met	Gly	Thr	Gln	Lys 350	Phe	Lys
45	Аla	Ala	11e 355	Asp	Gly	Asn	Gly	Phe 360	Lys	Gly	Thr	Тгр	Thr 365	Glu	Asn	Gly
	Gly	Gly 370	Asp	Va1	Ser	Gly	Arg 375	Phe	Туг	Gly	Pro	Ala 380	Gly	Glu	Glu	٧a٦
50	Ala 385	Gly	Lys	Tyr	Ser	Tyr 390	Arg	Pro	Thr	Asp	Ala 395	Glu	Lys	Gly	Gly	Phe 400
	Gly	۷al	Phe	Ala	Gly 405	Lys	Lys	Glu	Gln	Asp 410						
55	<21 <21 <21	1>	81 9 PRT													

```
<213>
                      Artificial Sequence
             <220>
              <223>
                      920L N-terminal
             <400>
             His Arg Val Trp Val Glu Thr Ala His
              <210>
10
              <211>
                      16
              <212>
                      PRT
              <213>
                      Artificial Sequence
              <220>
              <223>
                      953L N-terminal
15
              <400>
             Ala Thr Tyr Lys Val Asp Glu Tyr His Ala Asn Ala Arg Phe Ala Phe 1 5 10 15
              <210>
                      83
20
              <211>
                      9
              <212>
                      Artificial Sequence
              <213>
              <220>
                      519.1L N-terminal
              <223>
25
              <400>
             Met Glu Phe Phe Ile Ile Leu Leu Ala
1 5
              <210>
                      84
                      488
              <211>
30
              <212>
                      PRT
                      Artificial Sequence
              <213>
              <220>
                      deltaG287
              <223>
35
             Met Phe Lys Arg Ser Val Ile Ala Met Ala Cys Ile Phe Ala Leu Ser
1 5 10 15
             Ala Cys Gly Gly Gly Gly Gly Ser Pro Asp Val Lys Ser Ala Asp 20 25 30
40
             Thr Leu Ser Lys Pro Ala Ala Pro Val Val Ser Glu Lys Glu Thr Glu 35 40 45
             Ala Lys Glu Asp Ala Pro Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro 50 60
45
             Ser Ala Gln Gly Ser Gln Asp Met Ala Ala Val Ser Glu Glu Asn Thr 65 70 75 80
             Gly Asn Gly Gly Ala Val Thr Ala Asp Asn Pro Lys Asn Glu Asp Glu
85 90 95
50
             Val Ala Gln Asn Asp Met Pro Gln Asn Ala Gly Thr Asp Ser Ser 100 105 110
             Thr Pro Asn His Thr Pro Asp Pro Asn Met Leu Ala Gly Asn Met Glu 115 120 125
55
             Asn Gln Ala Thr Asp Ala Gly Glu Ser Ser Gln Pro Ala Asn Gln Pro
130 140
```

Asp Met Ala Asn Ala Ala Asp Gly Met Gln Gly Asp Asp Pro Ser Ala 145 150 160 Gly Gly Gln Asn Ala Gly Asn Thr Ala Ala Gln Gly Ala Asn Gln Ala 165 170 175 Gly Asn Asn Gln Ala Ala Gly Ser Ser Asp Pro Ile Pro Ala Ser Asn 180 185 190Pro Ala Pro Ala Asn Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala 195 200 205 10 Asn Gly Val Leu Ile Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His 210 215 220 Cys Lys Gly Asp Ser Cys Ser Gly Asn Asn Phe Leu Asp Glu Glu Val 225 230 240 15 Gln Leu Lys Ser Glu Phe Glu Lys Leu Ser Asp Ala Asp Lys Ile Ser 245 250 255 Asn Tyr Lys Lys Asp Gly Lys Asn Asp Lys Phe Val Gly Leu Val Ala 260 265 270 20 Asp Ser Val Gln Met Lys Gly Ile Asn Gln Tyr Ile Ile Phe Tyr Lys 275 280 285 Pro Lys Pro Thr Ser Phe Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg 290 295 300 25 Arg Ser Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp 305 310 315 320 Thr Leu Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly 325 330 33530 Asn Ile Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala 340 345 350 Glu Lys Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro 355 360 365 35 Ala Lys Gly Glu Met Leu Ala Gly Ala Ala Val Tyr Asn Gly Glu Val 370 375 380 Leu His Phe His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg Gly Arg 385 390 395 400 40 Phe Ala Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile 405 410 415Asp Ser Gly Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala 420 43045 Ile Asp Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Ser Gly 435 440 445 Asp Val Ser Gly Lys Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly 450 460 50 Lys Tyr Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val 465 470 475 Phe Ala Gly Lys Lys Glu Gln Asp 485 55 <210> 85 712 <211>

	<212 <213		PRT Artificial Sequence													
5		<220> <223>		2												
	<400 Met 1		85 Asn	Pro	Leu 5	val	Asn	Gln	Ala	Ala 10	Met	۷al	Leu	Pro	۷a٦ 15	Phe
10	Leu	Leu	Ser	Ala 20	Cys	Leu	Gly	Gly	G]y 25	Gly	Ser	Phe	Asp	Leu 30	Asp	Ser
	۷al	Asp	Thr 35	Glu	Ala	Pro	Arg	Pro 40	Ala	Pro	Lys	туг	G]n 45	Asp	۷a٦	Phe
15	Ser	Glu 50	Lys	Pro	Gln	Ala	G]n 55	Lys	Asp	Gln	Gly	G]y 60	Tyr	Gly	Phe	Аlа
	Met 65	Arg	Leu	Lys	Arg	Arg 70	Asn	Тгр	туг	Pro	G]n 75	Аlа	Lys	Glu	Asp	Glu 80
20	٧a٦	Lys	Leu	Asp	G]u 85	Ser	Asp	Trp	Glu	А1а 90	Thr	Gly	Leu	Pro	Asp 95	Glu
	Pro	Lys	Glu	Leu 100	Pro	Lys	Arg	Gln	Lys 105	Ser	٧a٦	Ile	Glu	Lys 110	٧a٦	Glu
25	Thr	Asp	Ser 115	Asp	Asn	Asn	Ile	Tyr 120	Ser	Ser	Pro	Tyr	Leu 125	Lys	Pro	Ser
	Asn	ніs 130	Gln	Asn	Gly	Asn	Thr 135	Glу	Asn	Gly	Ile	Asn 140	Gln	Pro	Lys	Asn
30	Gln 145	Ala	Lys	Asp	Tyr	Glu 150	Asn	Phe	Lys	Tyr	Val 155	Tyr	Ser	Gly	Trp	Phe 160
	Tyr	Lys	His	Ala	Lys 165	Arg	Glu	Phe	Asn	Leu 170	Lys	val	Glu	Pro	Lys 175	Ser
35	Ala	Lys	Asn	Gly 180	Asp	Asp	Gly	Tyr	11e 185	Phe	Tyr	His	Gly	Lys 190	Glu	Pro
	Ser	Arg	G]n 195	Leu	Pro	Ala	Ser	G]y 200	Lys	Ile	Thr	Tyr	Lys 205	Gly	٧a٦	Trp
40	His	Phe 210	Ala	Thr	Asp	Thr	Lys 215	Lys	Gly	Gln	Lys	Phe 220	Arg	Glu	Ile	Ile
	Gln 225	Pro	Ser	Lys	Ser	G]n 230	Gly	Asp	Arg	Tyr	Ser 235	Gly	Phe	Ser	Gly	Asp 240
45	Asp	Gly	Glu	Glu	Tyr 245	Ser	Asn	Lys	Asn	Lys 250	Ser	Thr	Leu	Thr	Asp 255	Glу
.0	Gln	Glu	Gly	Туг 260	Gly	Phe	Thr	Ser	Asn 265	Leu	Glu	val	Asp	Phe 270	His	Asn
50	Lys	Lys	Leu 275	Thr	Gly	Lys	Leu	11e 280	Arg	Asn	Asn	Ala	Asn 285	Thr	Asp	Asn
50	Asn	G1n 290	Ala	Thr	Thr	Thr	G1n 295	Tyr	Tyr	Ser	Leu	G1u 300	Ala	Gln	۷a٦	Thr
	Gly 305	Asn	Arg	Phe	Asn	Gly 310	Lys	Ala	Thr	Ala	Thr 315	Asp	Lys	Pro	Gln	G1n 320
55	Asn	Ser	Glu	Thr	Lys 325	Glu	His	Pro	Phe	Val 330	Ser	Asp	Ser	Ser	Ser 335	Leu

Ser Gly Gly Phe Phe Gly Pro Gln Gly Glu Glu Leu Gly Phe Arg Phe 340 345 350Leu Ser Asp Asp Gln Lys Val Ala Val Gly Ser Ala Lys Thr Lys 355 360 365 Asp Lys Pro Ala Asn Gly Asn Thr Ala Ala Ala Ser Gly Gly Thr Asp 370 375 380 Ala Ala Ala Ser Asn Gly Ala Ala Gly Thr Ser Ser Glu Asn Gly Lys 385 390 395 400 10 Leu Thr Thr Val Leu Asp Ala Val Glu Leu Lys Leu Gly Asp Lys Glu 410 415 Val Gln Lys Leu Asp Asn Phe Ser Asn Ala Ala Gln Leu Val Val Asp 420 425 430 15 Gly Ile Met Ile Pro Leu Leu Pro Glu Ala Ser Glu Ser Gly Asn Asn 435 440 445 Gln Ala Asn Gln Gly Thr Asn Gly Gly Thr Ala Phe Thr Arg Lys Phe 450 455 460 20 Asp His Thr Pro Glu Ser Asp Lys Lys Asp Ala Gln Ala Gly Thr Gln 465 470 480 Thr Asn Gly Ala Gln Thr Ala Ser Asn Thr Ala Gly Asp Thr Asn Gly 485 490 495 25 Lys Thr Lys Thr Tyr Glu Val Glu Val Cys Cys Ser Asn Leu Asn Tyr 500 505 510 Leu Lys Tyr Gly Met Leu Thr Arg Lys Asn Ser Lys Ser Ala Met Gln 515 525 30 Ala Gly Glu Ser Ser Gln Ala Asp Ala Lys Thr Glu Gln Val Glu 530 540 Gln Ser Met Phe Leu Gln Gly Glu Arg Thr Asp Glu Lys Glu Ile Pro 545 550 555 35 Ser Glu Gln Asn Ile Val Tyr Arg Gly Ser Trp Tyr Gly Tyr Ile Ala 565 570 575 Asn Asp Lys Ser Thr Ser Trp Ser Gly Asn Ala Ser Asn Ala Thr Ser 580 590 40 Gly Asn Arg Ala Glu Phe Thr Val Asn Phe Ala Asp Lys Lys Ile Thr 595 600 605 Gly Thr Leu Thr Ala Asp Asn Arg Gln Glu Ala Thr Phe Thr Ile Asp 610 620 45 Gly Asn Ile Lys Asp Asn Gly Phe Glu Gly Thr Ala Lys Thr Ala Glu 625 630 635 Ser Gly Phe Asp Leu Asp Gln Ser Asn Thr Thr Arg Thr Pro Lys Ala 645 650 655 50 Tyr Ile Thr Asp Ala Lys Val Gln Gly Gly Phe Tyr Gly Pro Lys Ala 660 670 Glu Glu Leu Gly Gly Trp Phe Ala Tyr Pro Gly Asp Lys Gln Thr Lys 675 680 685 55 Asn Ala Thr Asn Ala Ser Gly Asn Ser Ser Ala Thr Val Val Phe Gly 690 700

	Ala 705	Lys	Arg	Gln	Gln	Pro 710	۷al	Arg								
5	<210 <211 <211 <211	1> 2>	86 274 PRT Art	ific <sup>-</sup>	ial s	Seque	ence									
	<220> <223>		741													
10	<400 val 1	0> Asn	86 Arg	Thr	<b>а</b> 1а 5	Phe	Cys	Cys	Leu	Ser 10	Leu	Thr	Thr	Аlа	Leu 15	Ile
15	Leu	Thr	Ala	Cys 20	Ser	Ser	Gly	Gly	Gly 25	Gly	val	Ala	Ala	Asp 30	Ile	Gly
	Аla	Gly	Leu 35	Ala	Asp	Ala	Leu	Thr 40	Ala	Pro	Leu	Asp	His 45	Lys	Asp	Lys
20	Gly	Leu 50	Gln	Ser	Leu	Thr	Leu 55	Asp	Gln	Ser	Val	Arg 60	Lys	Asn	Glu	Lys
	Leu 65	Lys	Leu	Ala	Ala	G]n 70	Gly	Ala	Glu	Lys	Thr 75	Tyr	Gly	Asn	Glу	Asp 80
<i>25 30</i>	Ser	Leu	Asn	Thr	G]y 85	Lys	Leu	Lys	Asn	Asp 90	Lys	Val	Ser	Arg	Phe 95	Asp
	Phe	Ile	Arg	Gln 100	Ile	G∂u	Val	Asp	Gly 105	Gln	Leu	Ile	Thr	Leu 110	Glu	Ser
	Gly	Glu	Phe 115	Gln	val	Tyr	Lys	Gln 120	Ser	нis	ser	Ala	Leu 125	Thr	Ala	Phe
	Gln	Thr 130	Glu	Gln	Ile	G∏n	Asp 135	Ser	Glu	His	Ser	Gly 140	Lys	Met	val	Ala
35	Lys 145	Arg	Gln	Phe	Arg	17e 150	Gly	Asp	Ile	Ala	Gly 155	Glu	His	Thr	Ser	Phe 160
	Asp	Lys	Leu	Pro	Glu 165	Gly	Gly	Arg	Ala	Thr 170	Tyr	Arg	GТу	Thr	Ala 175	Phe
40	Gly	Ser	Asp	Asp 180	Ala	Gly	Gly	Lys	Leu 185	Thr	Tyr	Thr	Ile	Asp 190	Phe	Ala
40	ΑΊа	Lys	Gln 195		Asn	Gly		11e 200		His	Leu	Lys	Ser 205	Pro	Glu	Leu
45	Asn	Val 210	Asp	Leu	Ala	Ala	Ala 215	Asp	Ile	Lys	Pro	Asp 220	Gly	Lys	Arg	His
45	Ala 225	٧al	Ile	Ser	Gly	Ser 230	۷al	Leu	Tyr	Asn	G]n 235	Ala	Glu	Lys	Gly	Ser 240
	Tyr	Ser	Leu	Gly	Ile 245	Phe	Gly	Gly	Lys	Ala 250	Gln	Glu	val	Ala	G]y 255	Ser
50	Ala	Glu	٧a٦	Lys 260	Thr	٧al	Asn	Gly	11e 265	Arg	His	Ile	Gly	Leu 270	Ala	Ala
	Lys	Gln														
55	<210 <210		87 1082	2												

<212> **PRT** <213> Artificial Sequence <220> <223> 983 <400> Met Arg Thr Thr Pro Thr Phe Pro Thr Lys Thr Phe Lys Pro Thr Ala 1 5 10 15 Met Ala Leu Ala Val Ala Thr Thr Leu Ser Ala Cys Leu Gly Gly 25 30 10 Gly Gly Gly Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile 35 40 Gly Ser Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr 50 60 Ala Gly Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala 70 75 80Gly Arg Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala 90 95 20 Pro Pro Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala  $100 \hspace{1.5cm} 105 \hspace{1.5cm} 110$ Tyr Lys Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr 115 120 125 Gly Arg Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly 130 140 Ser Ile Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn 145 150 160 30 Glu Asn Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu 165 170 Asp Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val 180 185 35 Ile Glu Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile 195 200 205 Gly His Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp 210 220 40 Gly Arg Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met 225 230 235 Asn Thr Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg 245 250 255 45 Asn Ala Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn 260 265 270 Ser Phe Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile 275 280 285 50 Ala Asn Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly 290 295 Gly Asp Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr 305 315 32055 Gly Asn Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe 325 330

Ser Thr Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu 340 350Pro Phe Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly 355 360 365 Val Asp Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro 370 380 Gly Thr Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala 385 390 400 10 Met Trp Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg 405 410 415 Thr Asn Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val 420 425 43015 Thr Gly Thr Ala Ala Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn 435 Asp Asn Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala 450 460 20 Val Gly Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys 465 470 475 Ala Met Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp 485 490 495 25 Thr Lys Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser 500 505 Gly Thr Gly Gly Leu Ile Lys Lys Gly Gly Ser Gln Leu Gln Leu His 515 520 525 30 Gly Asn Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly Ser Leu 530 540 Val Leu Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly 545 550 560 35 Ala Leu Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp 565 570 Gly Ile Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr 580 585 590 40 Val His Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr 595 600 Thr Arg Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly 610 620 45 Gly Lys Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn 625 630 635 Ser Thr Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln 645 650 655 50 Asp Tyr Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala 660 670 Ser Leu Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu 675 680 685 Ser Tyr Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala 690 695 700 55

	Ala 705	His	Ser	Ala	Pro	Ala 710	Gly	Leu	Lys	His	Ala 715	Val	Glu	G∏n	Gly	Gly 720
5	Ser	Asn	Leu	Glu	Asn 725	Leu	Met	val	Glu	Leu 730	Asp	Ala	Ser	Glu	Ser 735	Ser
	Ala	Thr	Pro	G1u 740	Thr	٧al	Glu	Thr	Ala 745	Ala	Ala	Asp	Arg	Thr 750	Asp	Met
10	Pro	Gly	11e 755	Arg	Pro	Tyr	Gly	Ala 760	Thr	Phe	Arg	Ala	Ala 765	Ala	Аlа	Val
	Gln	His 770	Ala	Asn	Ala	Ala	Asp 775	Gly	val	Arg	Ile	Phe 780	Asn	Ser	Leu	Ala
15	Ala 785	Thr	٧al	Tyr	Ala	Asp 790	Ser	Thr	Ala	Ala	His 795	Ala	Asp	Met	Gln	Gly 800
	Arg	Arg	Leu	Lys	Ala 805	val	Ser	Asp	Glу	Leu 810	Asp	His	Asn	GТу	Thr 815	Gly
20	Leu	Arg	val	11e 820	Ala	Gln	Thr	Gln	G]n 825	Asp	Gly	Gly	Thr	Trp 830	Glu	Gln
	Gly	Gly	va1 835	Glu	Gly	Ly5	Met	Arg 840	Gly	Ser	Thr	Gln	Thr 845	۷al	Gly	Ile
25	Аlа	Ala 850	Lys	Thr	Gly	Glu	Asn 855	Thr	Thr	Ala	Ala	А1а 860	Thr	Leu	Gly	Met
	Gly 865	Arg	Ser	Thr	Trp	Ser 870	Glu	Asn	Ser	Ala	Asn 875	Ala	Lys	Thr	Asp	Ser 880
30	Ile	Ser	Leu	Phe	Ala 885	Gly	Ile	Arg	His	Asp 890	Ala	Gly	Asp	Ile	G]y 895	Tyr
	Leu	Lys	Gly	Leu 900	Phe	Ser	Tyr	Gly	Arg 905	Tyr	Lys	Asn	Ser	Ile 910	Ser	Arg
35	Ser	Thr	Gly 915	Ala	Asp	Glu	His	Ala 920	Glu	Gly	Ser	٧al	Asn 925	Gly	Thr	Leu
	Met	G]n 930	Leu	Gly	Ala	Leu	G]y 935	Gly	val	Asn	val	Pro 940	Phe	Ala	ΑΊа	Thr
40	Gly 945	Asp	Leu	Thr	val	G1u 950	Gly	Gly	Leu	Arg	Tyr 955	Asp	Leu	Leu	Lys	Gln 960
	Asp	Ala	Phe	Ala	Glu 965	Lys	Gly	Ser	Ala	Leu 970	Gly	Trp	Ser	Gly	Asn 975	Ser
45	Leu	Thr	Glu	Gly 980	Thr	Leu	val	Gly	Leu 985	Ala	Gly	Leu	Lys	Leu 990	Ser	Gln
	Pro	Leu	Ser 995	Asp	Lys	Ala	val	Leu 1000		Ala	Thr	Ala	Gly 100		Glu	Arg
50	Asp	Leu 1010		Gly	Arg	Asp	Tyr 1015		۷al	Thr	Gly	Gly 1020		Thr	Gly	Ala
50	⊤hr 102!		Ala	Thr	Gly	Lys 1030	⊤hr )	Gly	Ala	Arg	Asn 103		Pro	His	Thr	Arg 1040
	Leu	val	Ala	Gly	Leu 1045		Ala	Asp	٧a٦	Glu 1050		Gly	Asn	Gly	Trp 105	
55	Gly	Leu	Ala	Arg 1060		Ser	Tyr	Ala	Gly 106		Lys	Gln	Tyr	Gly 1070		His

```
Ser Gly Arg Val Gly Val Gly Tyr Arg Phe
1075 1080
                     <210>
                                         88
2505
                     <211>
                     <212>
                                        DNA
                     <213>
                                        Artificial Sequence
                     <220>
                     <223>
                                        deltaG287-919
10
                     <400>
                                         88
                     atggctagcc ccgatgttaa atcggcggac acgctgtcaa aaccggccgc tcctgttgtt
                                                                                                                                                                                               120
                     gctgaaaaag agacagaggt aaaagaagat gcgccacagg caggttctca aggacagggc
                     gcgccatcca cacaaggcag ccaagatatg gcggcagttt cggcagaaaa tacaggcaat
ggcggtgcgg caacaacgga caaacccaaa aatgaagacg agggaccgca aaatgatatg
                                                                                                                                                                                               180
                                                                                                                                                                                               240
                     ccgcaaaatt ccgccgaatc cgcaaatcaa acagggaaca accaacccgc cgattcttca
15
                                                                                                                                                                                                300
                    gattccgcc ccgcgtcaaa ccctgcacct gcgaatggcg gtagcaattt tggaagggtt gatttggcta atggcgtttt gattgatggg ccgtcgcaaa atataacgtt gacccactgt aaaggcgatt cttgtaatgg tgataattta ttggatgaag aagcaccgtc aaaatcagaa ttgaaaatt taaatgagtc tgaacgaatt gagaaatata agaaagatgg gaaaagcgat aaatttacta attggttgc gacagcagtt caagctaatg gaactaacaa atatgtcatc attataaag acaagtccgc ttcatcttca tctgcgcgat tcaggcgtt tgcacggtcg
                                                                                                                                                                                               360
                                                                                                                                                                                               420
                                                                                                                                                                                               480
                                                                                                                                                                                               540
                                                                                                                                                                                               600
                                                                                                                                                                                               660
20
                     aggaggtege tteetgeega gatgeegeta ateeegtea ateaggegga täegetgatt
                                                                                                                                                                                               720
                    gtcgatgggg aagcggtcag cctgacgggg cattccggca atatcttcgc gcccgaaggg aattaccggt atctgacta cggggcggaa aaattgcccg gcggatcgta tgccctccgt gtgcaaggcg aaccggcaaa aggcgaaatg cttgctgga cggccgtgta caacggcgaa
                                                                                                                                                                                               780
                                                                                                                                                                                               840
                                                                                                                                                                                               900
                    gtgcaaggcg aaccggcaaa aggcgaaaty cligitygta cyyclyta caacyycyaa gtgctgcatt ttcatacgga aaacggccgt ccgtacccga ctagaggcag gtttgccgca aaagtcgatt tcggcagcaa atctgtggac ggcattatcg acagcggcga tgatttgcat atgggtacgc aaaaattcaa agccgccatc gatggaaacg gcttaaggg gacttggacg gaaaatggcg gcggggagtg ttccggaagg ttttacggcc cggccggcga ggaagtgggc ggaaaataca gctatcgccc gacagatgcg gaaaagggcg gattcggcgt gtttgccggc aaaaaagagc aggatggatc cggaggagga ggatgccaaa gcaagagcat ccaaaccttt
                                                                                                                                                                                               960
                                                                                                                                                                                               1020
25
                                                                                                                                                                                               1080
                                                                                                                                                                                               1140
                                                                                                                                                                                               1200
                                                                                                                                                                                               1260
                     ccgcaacccg acacatccgt catcaacggc ccggaccggc cggtcggcat ccccgacccc
                                                                                                                                                                                               1320
                     gccggaacga cggtcggcgg cggcggggcc gtctataccg ttgtaccgca cctgtccctg
ccccactggg cggcgcagga tttcgccaaa agcctgcaat ccttccgcct cggctgcgcc
                                                                                                                                                                                               1380
30
                                                                                                                                                                                               1440
                    aatttgaaaa accgccaagg ctggcaggat gtgtgcgcc aagcctttca aacccccgtc cattccttc aggcaaaaca gtttttgaa cgctattca cgccgtggca ggttgcaggc aacggaagcc ttgccggtac ggttaccggc tattacgagc cggtgctgaa gggcgacgac aggcggacgg cacaagcccg cttcccgatt tacggtattc ccgacgattt tatctccgtc
                                                                                                                                                                                               1500
                                                                                                                                                                                               1560
                                                                                                                                                                                               1620
                                                                                                                                                                                               1680
                    aggcggacgg cacaagcccg cttcccgatt tacggtattc ccgacgattt tatctccgtc cccctgcctg ccggtttgcg gagcggaaaa gcccttgtcc gcatcaggca gacggggaaaa acaagcggca caatcgacaa taccggcgc acacataccg ccgacctctc ccgattccc atcaccgcgc gcacaacggc aatcaaaggc aggtttgaag gaagccgctt cctccctac cacacgcgca accaaatcaa cggcggcgcg cttgacggca aagccccgat actcggttac gccgaagacc ccgtcggaact ttttttatg cacatccaag gctcgggccg tctgaaaacc ccgtccggca aatacatccg catcggctat gccgacaaaa acgaacatcc ctacgtttcc atcggacgct atatggcgga caaaggctac ctcaagctcg ggcagacctc gatgcagggc atcaaagcct atatgcggca aaatccgcaa cgcctcgccg aagttttggg tcaaaacccc agcacaccct tgatgagga atatgccgga aggagcacac aggagcactt tgatgagga atatgccgga gcagcaccc ggcactacat taccttggg
                                                                                                                                                                                               1740
35
                                                                                                                                                                                               1800
                                                                                                                                                                                               1860
                                                                                                                                                                                               1920
                                                                                                                                                                                               1980
                                                                                                                                                                                               2040
                                                                                                                                                                                               2100
                                                                                                                                                                                               2160
2220
40
                    ggcacgccgt tgatggggga atatgccggc gcagtcgacc ggcactacat taccttgggc gcgcccttat ttgtcgccac cgcccatccg gttacccgca aagccctcaa ccgcctgatt atggcgcagg ataccggcag cgcgattaaa ggcgcggtgc gcgtggatta tttttgggga tacggcgacg aagccggca acttgccggc aaacaggaaca ccacgggtta cgtctggcag
                                                                                                                                                                                               2280
                                                                                                                                                                                               2340
2400
                                                                                                                                                                                               2460
                     ctcctaccca acggtatgaa gcccgaatac cgcccgtaac tcgag
                                                                                                                                                                                               2505
45
                     <210>
                                         89
                     <211>
                                        832
                     <212>
                     <213>
                                        Artificial Sequence
50
                     <220>
                     <223>
                                        deltaG287-919
                     <400>
                    Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala
1 10 15
55
                    Ala Pro Val Val Ala Glu Lys Glu Thr Glu Val Lys Glu Asp Ala Pro
```

Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Thr Gln Gly Ser Gln 35 40 Asp Met Ala Ala Val Ser Ala Glu Asn Thr Gly Asn Gly Gly Ala Ala 50 60 Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Pro Gln Asn Asp Met 65 70 75 10 Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln Thr Gly Asn Asn Gln Pro Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser Asn Pro Ala Pro Ala Asn 100 105 110 15 Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala Asn Gly Val Leu Ile 115 120 125Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser 130 135 140 20 Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu Ala Pro Ser Lys Ser Glu 145 150 160 Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile Glu Lys Tyr Lys Lys Asp 165 170 175Gly Lys Ser Asp Lys Phe Thr Asn Leu Val Ala Thr Ala Val Gln Ala  $180 \hspace{1cm} 185 \hspace{1cm} 190$ Asn Gly Thr Asn Lys Tyr val Ile Ile Tyr Lys Asp Lys Ser Ala Ser 195 200 205 Ser Ser Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser Leu 210 220 30 Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu Ile 225 230 235 240 Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile Phe 245 250 35 Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys Leu 260 265 270 Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ala Lys Gly 275 280 285 40 Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His Phe 290 295 300 His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg Gly Arg Phe Ala Ala 305 310 315 45 Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser Gly 325 330 335 Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp Gly 340 34550 Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Gly Asp Val Ser 355 360 365 Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly Lys Tyr Ser 55 Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val Phe Ala Gly

	385					390					395					400
5	Lys L	ys G	šlu (	Gln	Asp 405	Gly	ser	Gly	Gly	Gly 410	Gly	Cys	Gln	ser	Lys 415	Ser
3	Ile G	Gln T		Phe 420	Pro	Gln	Pro	Asp	Thr 425	Ser	۷al	Ile	Asn	Gly 430	Pro	Asp
	Arg F		/al (	Glу	ıle	Pro	Asp	Pro 440	Аlа	Glу	Thr	Thr	va1 445	GТу	Glу	Gly
10	Gly A	Ala V 450	/al ·	Tyr	Thr	val	va1 455	Pro	His	Leu	Ser	Leu 460	Pro	His	Тгр	Аlа
15	Ala 6 465	Gln A	Asp	Phe	Ala	Lys 470	Ser	Leu	Gln	Ser	Phe 475	Arg	Leu	Gly	Cys	Ala 480
	Asn L	_eu L	_ys /	Asn	Arg 485	Gln	Gly	Trp	Gln	Asp 490	val	Cys	Аlа	Gln	Ala 495	Phe
	Gln T	Thr P		val 500	His	Ser	Phe	Gln	Ala 505	Lys	Gln	Phe	Phe	Glu 510	Arg	Tyr
20	Phe T		Pro 5	тгр	Gln	val	Аlа	G]y 520	Asn	Glу	Ser	Leu	Ala 525	Gly	Thr	۷al
	Thr G	Gly T 530	Гуr	Tyr	Glu	Pro	va1 535	Leu	Lys	Gly	Asp	Asp 540	Arg	Arg	Thr	Ala
25	Gln A 545	Ala A	Arg	Phe	Pro	11e 550	Tyr	Gly	Ile	Pro	Asp 555	Asp	Phe	Ile	Ser	val 560
	Pro L	_eu F	Pro .	Ala	G]y 565	Leu	Arg	Ser	Gly	Lys 570	ΑΊа	Leu	٧a٦	Arg	11e 575	Arg
30	Gln 1	Thr G		Lys 580	Asn	Ser	Gly	Thr	11e 585	Asp	Asn	Thr	Gly	G]y 590	Thr	His
	Thr A		Asp 595	Leu	Ser	Arg	Phe	Pro 600	Ile	Thr	Ala	Arg	Thr 605	Thr	Ala	Ile
35	Lys 6	Gly A 510	arg	Phe	Glu	GТу	Ser 615	Arg	Phe	Leu	Pro	Tyr 620	His	Thr	Arg	Asn
	Gln 1 625	Ile A	Asn	G∃y	Gly	Ala 630	Leu	Asp	Gly	Lys	Ala 635	Pro	Ile	Leu	Gly	Tyr 640
40	Ala G	Glu A	Asp	Pro	va1 645	Glu	Leu	Phe	Phe	Met 650	His	Ile	G∏n	Gly	ser 655	Gly
	Arg L	_eu L		Thr 6 <b>6</b> 0	Pro	ser	Glу	Lys	Tyr 665	Ile	Arg	Ile	Gly	Tyr 670	Ala	Asp
45	Lys A		67u 675	нis	Pro	Tyr	٧a٦	Ser 680	Ile	Gly	Arg	Tyr	Met 685	Ala	Asp	Lys
	Gly T	Tyr L 590	_eu	Lys	Leu	Gly	G]n 695	Thr	Ser	Met	G∏n	G]y 700	Ile	Lys	Ala	Tyr
50	Met <i>A</i> 705	Arg G	Gln .	Asn	Pro	G]n 710	Arg	Leu	Ala	Glu	Val 715	Leu	Gly	Gln	Asn	Pro 720
	Ser 1	Tyr I	[]e	Phe	Phe 725	Arg	Glu	Leu	Ala	Gly 730	Ser	Ser	Asn	Asp	Gly 735	Pro
<i>55</i>	val (	Gly A		Leu 740	GÌу	Thr	Pro	Leu	Met 745	GТу	Glu	Tyr	Ala	Gly 750	Ala	val
	Asp A	Arg ⊦	His	Tyr	Ile	Thr	Leu	G∃y	Ala	Pro	Leu	Phe	val	Ala	Thr	Аlа

```
765
                                                             755
                                His Pro Val Thr Arg Lys Ala Leu Asn Arg Leu Ile Met Ala Gln Asp
770 780
                                Thr Gly Ser Ala Ile Lys Gly Ala Val Arg Val Asp Tyr Phe Trp Gly 785 790 795
                                Tyr Gly Asp Glu Ala Gly Glu Leu Ala Gly Lys Gln Lys Thr Thr Gly 805 810 815
10
                                Tyr Val Trp Gln Leu Leu Pro Asn Gly Met Lys Pro Glu Tyr Arg Pro 820 825 830
                                 <210>
                                                              90
15
                                 <211>
                                                              1746
                                 <212>
                                                             DNA
                                                             Artificial Sequence
                                 <213>
                                 <220>
                                                             deltaG287-953
                                 <223>
20
                                 <400>
                                                             90
                                 atggctagcc ccgatgttaa atcggcggac acgctgtcaa aaccggccgc tcctgttgtt
                                                                                                                                                                                                                                                                                         60
                                 gctgaaaaag agacagaggt aaaagaagat gcgccacagg caggttctca aggacagggc
                                                                                                                                                                                                                                                                                         120
                               gctgaaaaag agacagaggt aaaagaagat gcgccacagg caggttctca aggacagggc gcgccatcca cacaaggcag ccaagatatg gcggcagttt cggcagaaaa tacaaggcaat gcggtgcgg caacaacgga caaacccaaa aatgaagacg agggaccgca aaatgatatg ccgcaaaatt ccgccgaatc cgcaaatcaa acagggaaca accaaccgc cgattcttca gattccgccc ccgcgtcaaa ccctgcacct gcgaatggcg gtagcaattt tggaagggtt gatttggcta atggcgttt gattgatggg ccgtcgcaaa atataacgtt gacccactgt aaaggcgatt cttgtaatgg tgataattta ttggatgaag aagcaccgtc aaaatcagaa tttgaaaatt taaatgagtc tgaacgaatt gagaaatata agaaagatgg gaaaagcgat aaatttacta atttggttgc gacagcagtt caagctaatg gaactaacaa atatgtcatc atttataaag acaagtccgc ttcatcttca tctgcgcgat tcaggcgttc tgcacggtcgagagagagaggcgc ttcctoccga gatgccgcta atcaggcggat tacaggcggt tacaggtcgat tacaggcggt tacaggtcgat tacaggcggt tacaggtcgat tacaggcggt tacaggtcgat tacaggcggt tacaggtcgat tacaggcggat tacaggcggt tacaggcggt tacaggcggt atcaggcggat tacaggcggat tacaggcaggat tacaggcggat tacaggcaggat tacaggcggat tacaggcaggat tacaggcggat tacaggcggat tacaggcaggat tacaggat tacaggcaggat tacaggcaggat tacaggat 
                                                                                                                                                                                                                                                                                         180
                                                                                                                                                                                                                                                                                          240
                                                                                                                                                                                                                                                                                          300
25
                                                                                                                                                                                                                                                                                         360
                                                                                                                                                                                                                                                                                          420
                                                                                                                                                                                                                                                                                         480
                                                                                                                                                                                                                                                                                         540
                                                                                                                                                                                                                                                                                         600
                               atttataaag acaagtccgc ttcatcttca tctgcgcgat tcaggcgttc tgcacggtcg aggaggtcgc ttcctgccga gatgccgcta atccccgtca atcaggcgga tacgctgatt gtcgatgggg aagcggtcag cctgacgggg cattccggca atatcttcgc gcccgaaggg aattaccggt atctgactta cggggcggaa aaattgcccg gcggatcgta tgcctccgt gtgcaaggcg aaccggcaaa aggcgaaatg cttgctggca cggccgtgta caacggcgaa gtgctgcatt ttcatacgga aaacggccgt ccgtacccga ctagaggcag gtttgccgca aaagtcgatt tcggcagcaa atctgtggac ggcattatcg acagcggcga tgatttgcat atgggtacgc aaaaattcaa agccgccatc gatggaaacg gctttaaggg gacttggacg gaaaatggcg gcggggatgt ttccggaagg ttttacggcc cggccggcga ggaagtggcg ggaaaataca gctatcgcc gacagatgcg gaaaagggcg gattcggcgt gtttgccgc aaaaaaaggc aggatggat cggaaggagga ggagccacct acaaagtgga cgaatatcac ggcctgacg gttcgccat cgaccattc aacaccagca ccaacgtcg cggttttacggctaccgg gttcgccat cgaccattc aacaccagca acggtaaaat cgacatcacc ggtctgacg gttcgcca gttcgccaa gccaacgc acggtaaaaat cgacatcacc
                                                                                                                                                                                                                                                                                         660
                                                                                                                                                                                                                                                                                         720
30
                                                                                                                                                                                                                                                                                          780
                                                                                                                                                                                                                                                                                         840
                                                                                                                                                                                                                                                                                         900
                                                                                                                                                                                                                                                                                         960
                                                                                                                                                                                                                                                                                         1020
                                                                                                                                                                                                                                                                                         1080
35
                                                                                                                                                                                                                                                                                         1140
                                                                                                                                                                                                                                                                                          1200
                                                                                                                                                                                                                                                                                         1260
                                                                                                                                                                                                                                                                                         1320
                                ggtctgaccg gttccgtcga gttcgaccaa gcaaaacgcg acggtaaaat cgacatcacc atccccgttg ccaacctgca aagcggttcg caacacttta ccgaccacct gaaatcagcc gacatcttcg atgccgcca atatccggac atccgctttg tttccaccaa attcaacttc
                                                                                                                                                                                                                                                                                         1380
                                                                                                                                                                                                                                                                                         1440
40
                                                                                                                                                                                                                                                                                         1500
                                 aacggcaaaa aactggtttc cgttgacggc aacctgacca tgcacggcaa aaccgccccc
                                                                                                                                                                                                                                                                                          1560
                                gtcaaactca aagccgaaaa attcaactgc taccaaagcc cgatggcgaa aaccgaagtt
tgcggcggcg acttcagcac caccatcgac cgcaccaaat ggggcgtgga ctacctcgtt
                                                                                                                                                                                                                                                                                         1620
                                                                                                                                                                                                                                                                                         1680
                                                                                                                                                                                                                                                                                         1740
                                 aacgītggta tgaccaaaag cgtccgcatc gacatccaaa tcgaggcagc caaacaataa
                                                                                                                                                                                                                                                                                          1746
                                 ctcgag
45
                                                             91
579
                                 <210>
                                 <211>
                                 <212>
                                                             PRT
                                                             Artificial Sequence
                                 <213>
                                 <220>
50
                                 <223>
                                                             deltaG287-953
                                Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala
1 10 15
55
                                Ala Pro Val Val Ala Glu Lys Glu Thr Glu Val Lys Glu Asp Ala Pro
20 25 30
```

Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Thr Gln Gly Ser Gln 35 40 Asp Met Ala Ala Val Ser Ala Glu Asn Thr Gly Asn Gly Gly Ala Ala 50 55 60 Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Pro Gln Asn Asp Met 65 70 75 80Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln Thr Gly Asn Asn Gln Pro 85 90 95 10 Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser Asn Pro Ala Pro Ala Asn  $100 \hspace{1.5cm} 105 \hspace{1.5cm} 10$ Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala Asn Gly Val Leu Ile  $115 \,$  120  $\,$  125 15 Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser 130 140 Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu Ala Pro Ser Lys Ser Glu 145 150 155 160 20 Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile Glu Lys Tyr Lys Lys Asp 165 170 175 Gly Lys Ser Asp Lys Phe Thr Asn Leu Val Ala Thr Ala Val Gln Ala 180 185 25 Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr Lys Asp Lys Ser Ala Ser 195 200 205Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser Leu 210 220 30 Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu Ile 225 230 235 240 Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile Phe 245 250 255 35 Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys Leu 260 265 270 Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ala Lys Gly 275 280 285 40 Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His Phe 290 295 300 His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg Gly Arg Phe Ala Ala 305 310 320 45 Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser Gly 325 330 335 Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp Gly 340 34550 Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Gly Asp Val Ser 355 360 365 Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly Lys Tyr Ser 370 375 380Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val Phe Ala Gly 385 390 395 400

```
Lys Lys Glu Gln Asp Gly Ser Gly Gly Gly Gly Ala Thr Tyr Lys Val
405 416 415
           Asp Glu Tyr His Ala Asn Ala Arg Phe Ala Ile Asp His Phe Asn Thr 420 425 430
           Ser Thr Asn Val Gly Gly Phe Tyr Gly Leu Thr Gly Ser Val Glu Phe 435 440 445
           Asp Gln Ala Lys Arg Asp Gly Lys Ile Asp Ile Thr Ile Pro Val Ala
450 455 460
10
           Asn Leu Gln Ser Gly Ser Gln His Phe Thr Asp His Leu Lys Ser Ala
465 470 475 480
           Asp Ile Phe Asp Ala Ala Gln Tyr Pro Asp Ile Arg Phe Val Ser Thr
485 490 495
15
           Lys Phe Asn Phe Asn Gly Lys Lys Leu Val Ser Val Asp Gly Asn Leu 500 510
           Thr Met His Gly Lys Thr Ala Pro Val Lys Leu Lys Ala Glu Lys Phe
515 520 525
           Asn Cys Tyr Gln Ser Pro Met Ala Lys Thr Glu Val Cys Gly Gly Asp 530 540
           Phe Ser Thr Thr Ile Asp Arg Thr Lys Trp Gly Val Asp Tyr Leu Val 545 550 560
25
           Asn Val Gly Met Thr Lys Ser Val Arg Ile Asp Ile Gln Ile Glu Ala
565 570 575
            Ala Lys Gln
30
                        92
2388
            <210>
            <211>
            <212>
                        DNA
            <213>
                        Artificial Sequence
35
            <220>
            <223>
                        deltaG287-961
            <400>
                                                                                                                        60
            atggctagcc ccgatgttaa atcggcggac acgctgtcaa aaccggccgc tcctgttgtt
            gcīgaaaaag agacagaggt aaaagaagat gcgccacagg caggīīctca aggacagggc
                                                                                                                        120
           gcgccatcca cacaaggcag ccaagatatg gcggcagttt cggcagaaaa tacaggcaat
ggcggtgcgg caacaacgga caaacccaaa aatgaagacg agggaccgca aaatgatatg
                                                                                                                        180
40
                                                                                                                        240
            ccgcaaaatt ccgccgaatc cgcaaatcaa acagggaaca accaacccgc cgattcttca
                                                                                                                        300
           gattccgccc ccgcgtcaaa ccctgcacct gcgaatggcg gtagcaattt tggaagggtt gatttggcta atggcgtttt gattgatggg ccgtcgcaaa atataacgtt gaccactgt aaaggcgatt cttgtaatgg tgataattta ttggatgaag aagcaccgtc aaaatcagaa
                                                                                                                        360
                                                                                                                        420
480
           tttgaaaatt taaatgagte tgaacgaatt gagaaatata agaaagatgg gaaaagcgat aaatttacta atttggttge gacagcagtt caagctaatg gaactaacaa atatgtcate atttataaag acaagtccge ttcatcttca tctgcgcgat tcaggcgtte tgcacggtcg
                                                                                                                        540
45
                                                                                                                        600
                                                                                                                        660
            aggaggtegč ttecťgeega gatgeegeta atečeégťea ateággégga táegeťgatť
                                                                                                                        720
           gťčgaťgggg aagcggtcág čctgacgggg cattccggca atatčťtčgc gccčgaággg
aattaccggt atctgactta cggggcggaa aaattgcccg gcggatcgta tgccctccgt
                                                                                                                        780
                                                                                                                        840
                                                                                                                        900
            gtgcaaggcg aaccggcaaa aggcgaaatg cttgctggca cggccgtgta caacggcgaa
50
           gtgctaggtg aactggcaaa aggcgaaatg cttgctggca tggctgtgta caatggcgaa
gtgctgcatt ttcatacgga aaacggccgt ccgtacccga ctagaggcag gtttgccgca
aaagtcgatt tcggcagcaa atctgtggac ggcattatcg acagcggcga tgatttgcat
atgggtacgc aaaaattcaa agccgccatc gatggaaacg gctttaaggg gacttggacg
gaaaatggcg gcggggatgt ttccggaagg ttttacggcc cggccggcga ggaagtggcg
                                                                                                                        960
                                                                                                                        1020
                                                                                                                        1080
                                                                                                                        1140
           ggaaaataca gctatcgccc gacagatgcg gaaaagggcg gattcggcgt gtttgccggc
aaaaaagagc aggatggatc cggaggagga ggagccacaa acgacgacga tgttaaaaaa
gctgccactg tggccattgc tgctgcctac aacaatggcc aagaaatcaa cggtttcaaa
                                                                                                                        1200
                                                                                                                        1260
55
                                                                                                                        1320
                                                                                                                        1380
            gctggagagā cčātctacga cāttgatgaa gacggcācaa ttāccaaaaa agācgcaact
```

```
gcagccgatg ttgaagccga cgactttaaa ggtctgggtc tgaaaaaagt cgtgactaac
                                                                                                                       1440
           ctgaccaaaa ccgtcaatga aaacaaacaa aacgtcgatg ccaaagtaaa agctgcagaa
                                                                                                                       1500
          tctgaaatag aaaagttaac aaccaagtta gcagacactg atgccgcttt agcagatact gatgccgctc tggatgcaac caccaacgcc ttgaataaat tgggagaaaa tataacgaca
                                                                                                                       1560
                                                                                                                       1620
           tttgctgaag agactaagac aaatatcgta aaaattgatg aaaaattaga agccgtggct
                                                                                                                       1680
          gataccgtcg acaagcatgc cgaagcattc aacgatatcg ccgattcatt ggatgaacc
aacactaagg cagacgaagc cgtcaaaacc gccaatgaag ccaaacagac ggccgaagaa
accaaacaaa acgtcgatgc caaagtaaaa gctgcagaaa ctgcagcagg caaagccgaa
gctgccgctg gcacagctaa tactgcagcc gacaaggccg aagctgtcgc tgcaaaagtt
accgacatca aagctgatat cgctacgaac aaagataata ttgctaaaaa agcaaacagt
gccgacgtgt acaccagaga agagtctgac agcaaaattg tcagaattga tggcctgaac
                                                                                                                       1740
                                                                                                                       1800
                                                                                                                       1860
                                                                                                                       1920
1980
                                                                                                                       2040
10
          gctactaccg aaaaattgga cacacgcttg gcttctgctg aaaaatccat tgccgatcac gatactcgcc tgaacggttt ggataaaaca gtgtcagacc tgcgcaaaga aacccgccaa ggccttgcag aacaagccgc gctctccggt ctgttccaac cttacaacgt gggtcggttc
                                                                                                                       2100
                                                                                                                       2160
                                                                                                                       2220
2280
          aatgtaacgg ctgcagtcgg cggctacaaa tccgaatcgg cagtcgccat cggtaccggc
ttccgcttta ccgaaaactt tgccgccaaa gcaggcgtgg cagtcggcac ttcgtccggt
tcttccgcag cctaccatgt cggcgtcaat tacgagtggt aactcgag
                                                                                                                       2340
                                                                                                                       2388
15
                       93
793
           <210>
           <211>
           <212>
                       Artificial Sequence
           <213>
20
           <220>
           <223>
                       deltaG287-961
          Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala
1 5 10
25
          Ala Pro Val Val Ala Glu Lys Glu Thr Glu Val Lys Glu Asp Ala Pro
20 25 30
          Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Thr Gln Gly Ser Gln
35 40 45
30
          Asp Met Ala Ala Val Ser Ala Glu Asn Thr Gly Asn Gly Gly Ala Ala 50 60
          Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Pro Gln Asn Asp Met 65 70 75 80
          Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln Thr Gly Asn Asn Gln Pro 85 90 95
          Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser Asp Pro Ala Pro Ala Asp 100 \hspace{1.5cm} 105 \hspace{1.5cm} 110
          Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala Asn Gly Val Leu Ile
115 120 125
          Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser
130 135 140
          Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu Ala Pro Ser Lys Ser Glu 145 150 155 160
          Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile Glu Lys Tyr Lys Lys Asp
165 170 175
          Gly Lys Ser Asp Lys Phe Thr Asn Leu Val Ala Thr Ala Val Gln Ala
180 185 190
50
          Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr Lys Asp Lys Ser Ala Ser
195 200 205
          Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser Leu 210 220
```

Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu Ile 225 230 235 240 Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile Phe 245 250 255 Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys Leu 260 265 270 Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ala Lys Gly 275 280 285 10 Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His Phe 290 295 300 His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg Gly Arg Phe Ala Ala 305 310 315 32015 Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser Gly 325 330 335 Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp Gly 340 34520 Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Asp Val Ser 355 360 365 Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly Lys Tyr Ser 370 380 25 Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val Phe Ala Gly 385 390 400 Lys Lys Glu Gln Asp Gly Ser Gly Gly Gly Gly Ala Thr Asn Asp Asp 415 30 Asp Val Lys Lys Ala Ala Thr Val Ala Ile Ala Ala Ala Tyr Asn Asn 420 425 430 Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile 435 440 445 35 Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp Ala Thr Ala Ala Asp Val 450 455 460 Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu Lys Lys Val Val Thr Asn 465 470 475 40 Leu Thr Lys Thr Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val 485 490 495 Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp 500 510 Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala Ala Leu Asp Ala Thr Thr 515 520 525 Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu 530 540 Thr Lys Thr Asn Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala 545 550 560 50 Asp Thr Val Asp Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp Ser 565 570 575 Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala Asn 580 585 590 55

```
Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala Lys
595 600 605
            Val Lys Ala Ala Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Ala Gly 610 620
            Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys Val 625 635 640
            Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn Lys Asp Asn Ile Ala Lys
645 650 655
10
            Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg Glu Glu Ser Asp Ser Lys 660 665 670
            Phe Val Arg Ile Asp Gly Leu Asn Ala Thr Thr Glu Lys Leu Asp Thr 675 680 685
15
            Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala Asp His Asp Thr Arg Leu
690 695
            Asn Gly Leu Asp Lys Thr Val Ser Asp Leu Arg Lys Glu Thr Arg Gln 705 710 715
20
            Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu Phe Gln Pro Tyr Asn
725 730 735
            Val Gly Arg Phe Asn Val Thr Ala Ala Val Gly Gly Tyr Lys Ser Glu
740 745 750
25
            Ser Ala Val Ala Ile Gly Thr Gly Phe Arg Phe Thr Glu Asn Phe Ala
755 760 765
            Ala Lys Ala Gly Val Ala Val Gly Thr Ser Ser Gly Ser Ser Ala Ala
770 780
30
            Tyr His Val Gly Val Asn Tyr Glu Trp
785 790
            <210>
                        94
            <211>
<212>
                        2700
                        DNA
35
            <213>
                        Artificial Sequence
            <220>
            <223>
                        deltaG287NZ-919
            <400>
40
            atggctagcc ccgatgtcaa gtcggcggac acgctgtcaa aacctgccgc ccctgttgtt
                                                                                                                       60
           tctgaaaaaag agacagaggc aaaggaagat gcgccacagg caggttctca aggacagggc
gcgccatccg cacaaggcgg tcaagatatg gcggcggttt cggaagaaaa tacaggcaat
ggcggtgcgg cagcaacgga caaacccaaa aatgaagacg agggggcgca aaatgatatg
ccgcaaaaag ccgccgataaa agatagttg acacccaaaa acccccggc ttcgaatatg
                                                                                                                       120
                                                                                                                       180
                                                                                                                       240
                                                                                                                       300
           ccggccggaa atatggaaaa ccaagcaccg gatgccgggg aatcggagca gccggcaaac caaccggata tggcaaatac ggcggacgga atgcagggtg acgatccgtc ggcaggcggg gaaaatgccg gcaatacggc tgcccaaggt acaaatcaag ccgaaaacaa tcaaaccgcc ggttctcaaa atcctgcctc ttcaaccaat cctagcgcca cgaatagcgg tggtgatttt
                                                                                                                       360
                                                                                                                       420
45
                                                                                                                       480
                                                                                                                       540
            ggaaggacga acgtgggcaa ttctgttgtg attgacgggc cgtcgcaaaa tataacgttg
acccactgta aaggcgattc ttgtagtggc aataatttct tggatgaaga agtacagcta
                                                                                                                       600
                                                                                                                       660
            aaatcagaat ttgaaaaatt aagtgatgca gacaaaataa gtaattacaa gaaagatggg
                                                                                                                       720
            aagaatgacg ggaagaatga taaatttgtc ggtttggttg ccgatagtgt gcagatgaag
ggaatcaatc aatatattat cttttataaa cctaaaccca cttcatttgc gcgatttagg
                                                                                                                       780
50
                                                                                                                       840
            cgttctgcac ggtcgaggcg gtcgcttccg gccgagatgc cgctgattcc cgtcaatcag
gcggatacgc tgattgtcga tggggaagcg gtcagcctga cggggcattc cggcaatatc
                                                                                                                       900
                                                                                                                       960
           ttcgcgcccg aagggaatta ccggtatctg acttacgggg cggaaaaatt gcccggcgga
tcgtatgccc tccgtgttca aggcgaacct tcaaaaggcg aaatgctcgc gggcacggca
gtgtacaacg gcgaagtgct gcattttcaa acggaaatata tagacagat tatagacag
                                                                                                                       1020
                                                                                                                       1080
                                                                                                                       1140
55
                                                                                                                       1200
            ggcaggtttg ccgcaaaagt cgatttcggc agcaaatctg tggacggcat tatcgacagc
            ggcgatggtī tgčatatggg tacgcaaaaa ttcaaagccg ccatcgatgg aaacggctīt
                                                                                                                       1260
```

```
1320
            ggcgaggaaq tqqcqqqaaa atacagctat cgcccaacag atgcggaaaa gggcggattc
                                                                                                               1380
            ggcgtgtttg ccggcaaaaa agagcaggat ggatccggag gaggaggatg ccaaagcaag
agcatccaaa cctttccgca acccgacaca tccgtcatca acggcccgga ccggccggtc
                                                                                                               1440
                                                                                                               1500
            ggcatccccg accccgccgg aacgacggtc ggcggcggcg gggccgtcta taccgttgta
                                                                                                               1560
            ccgcacctgt ccctgccca ctgggcggcg caggatttcg ccaaaagcct gcaatccttc cgcctcggct gcgccaattt gaaaaaccgc caaggctggc aggatgtgtg cgcccaagcc tttcaaaccc ccgtccattc ctttcaggca aaacagtttt ttgaacgcta tttcacgccg
                                                                                                               1620
                                                                                                               1680
                                                                                                               1740
           tggcaggttg caggcaacgg aagccttgcc ggtacggtta ccggctatta cgagccggtg ctgaagggcg acgacaggcg gacggcacaa gcccgcttcc cgattaccg tattcccgac gatttatct ccgtcccct gcctgccggt ttgcggagcg gaaaagccct tgtccgcatc aggcagacgg gaaaaacag cggcacaatc gacaataccg gcggcacaca taccgccgac ctctcccgat tcccatcac cgcgcgcaca acggcaatca aaggcaggtt tgaaggaagc
                                                                                                               1800
                                                                                                               1860
10
                                                                                                               1920
                                                                                                               1980
                                                                                                               2040
            cgcttcctcc cctaccacac gcgcaaccaa atcaacggcg gcgcgcttga cggcaaagcc ccgatactcg gttacgccga agaccccgtc gaactttttt ttatgcacat ccaaggctcg ggccgtctga aaaccccgtc cggcaaatac atcgcatcg gctatgccga caaaaacgaa catccttacg tttccatcgg acgctatatg gcggacaaag gctactcaa gctcggcag
                                                                                                               2100
                                                                                                               2160
                                                                                                               2220
15
                                                                                                               2280
            acctcgatgc agggcatcaa agcctatatg cggcaaaatc cgcaacgcct cgccgaagtt
                                                                                                               2340
            ttgggťcaaa accccagcta tátctttttc cýcgagcttg ccggaagcag caatgacýgt
                                                                                                               2400
            cccgtcggcg cactgggcac gccgttgatg ggggaatatg ccggcgcagt cgaccggcac
                                                                                                               2460
            tacattacct tgggcgcgc cttatttgtc gccaccgccc atccggttac ccgcaaagcc ctcaaccgcc tgattatggc gcaggatacc ggcagcgcga ttaaaggcgc ggtgcgcgtg
                                                                                                               2520
                                                                                                               2580
            gattattītt ggggatacgg cgacgaagcc ggcgaactīg ccggcaaaca gaaaaccacg
                                                                                                               2640
20
                                                                                                               2700
            ggttacgtct ggcagctcct acccaacggt atgaagcccg aataccgccc gtaaaagctt
            <210>
            <211>
                        897
            <212>
                       PRT
                       Artificial Sequence
            <213>
25
            <220>
            <223>
                        deltaG287NZ-919
            <400>
            Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala
1 10 15
30
            Ala Pro Val Val Ser Glu Lys Glu Thr Glu Ala Lys Glu Asp Ala Pro
20 25 30
            Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Ala Gln Gly Gln Gln 45
35
            Asp Met Ala Ala Val Ser Glu Glu Asn Thr Gly Asn Gly Gly Ala Ala
50 55
            Ala Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Ala Gln Asn Asp Met 65 70 75 80
40
            Pro Gln Asn Ala Asp Thr Asp Ser Leu Thr Pro Asn His Thr Pro 85 90 95
            Ala Ser Asn Met Pro Ala Gly Asn Met Glu Asn Gln Ala Pro Asp Ala
100 105 110
45
            Gly Glu Ser Glu Gln Pro Ala Asn Gln Pro Asp Met Ala Asn Thr Ala
115 120 125
            Asp Gly Met Gln Gly Asp Asp Pro Ser Ala Gly Glu Asn Ala Gly 130 140
50
            Asn Thr Ala Ala Gln Gly Thr Asn Gln Ala Glu Asn Asn Gln Thr Ala
145 150 155 160
            Gly Ser Gln Asn Pro Ala Ser Ser Thr Asn Pro Ser Ala Thr Asn Ser
165 170 175
            Gly Gly Asp Phe Gly Arg Thr Asn Val Gly Asn Ser Val Val Ile Asp
180 185 190
55
```

Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser Cys 195 200 205 Ser Gly Asn Asn Phe Leu Asp Glu Glu Val Gln Leu Lys Ser Glu Phe 210 220 Glu Lys Leu Ser Asp Ala Asp Lys Ile Ser Asn Tyr Lys Lys Asp Gly 225 230 240 Lys Asn Asp Gly Lys Asn Asp Lys Phe Val Gly Leu Val Ala Asp Ser 245 250 255 10 Val Gln Met Lys Gly Ile Asn Gln Tyr Ile Ile Phe Tyr Lys Pro Lys 260 265 270 Pro Thr Ser Phe Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser 275 280 285 15 Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu 290 295 300 Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile 305 310 315 320 20 Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys 325 330 335 Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ser Lys  $340 \hspace{1cm} 345 \hspace{1cm} 350$ 25 Gly Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His 355 360 365 Phe His Thr Glu Asn Gly Arg Pro Ser Pro Ser Arg Gly Arg Phe Ala 370 380 30 Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser 385 390 395 Gly Asp Gly Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp 405 410 415 35 Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Gly Asp Val 420 425 430 Ser Gly Lys Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly Lys Tyr 435 440 445 40 Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val Phe Ala 450 455 460 Gly Lys Lys Glu Gln Asp Gly Ser Gly Gly Gly Gly Cys Gln Ser Lys 465 470 475 480 45 Ser Ile Gln Thr Phe Pro Gln Pro Asp Thr Ser Val Ile Asn Gly Pro 485 490 495 Asp Arg Pro Val Gly Ile Pro Asp Pro Ala Gly Thr Thr Val Gly 500 505 50 Gly Gly Ala Val Tyr Thr Val Val Pro His Leu Ser Leu Pro His Trp 515 520 525 Ala Ala Gln Asp Phe Ala Lys Ser Leu Gln Ser Phe Arg Leu Gly Cys 530 540 Ala Asn Leu Lys Asn Arg Gln Gly Trp Gln Asp Val Cys Ala Gln Ala 545 550 555

Phe Gln Thr Pro Val His Ser Phe Gln Ala Lys Gln Phe Phe Glu Arg 565 570 575 Tyr Phe Thr Pro Trp Gln Val Ala Gly Asn Gly Ser Leu Ala Gly Thr 580 585 590 Val Thr Gly Tyr Tyr Glu Pro Val Leu Lys Gly Asp Asp Arg Arg Thr 595 600 605 Ala Gln Ala Arg Phe Pro Ile Tyr Gly Ile Pro Asp Asp Phe Ile Ser 610 615 620 10 Val Pro Leu Pro Ala Gly Leu Arg Ser Gly Lys Ala Leu Val Arg Ile 625 630 635 640 Arg Gln Thr Gly Lys Asn Ser Gly Thr Ile Asp Asn Thr Gly Gly Thr 645 650 655 15 His Thr Ala Asp Leu Ser Arg Phe Pro Ile Thr Ala Arg Thr Thr Ala 660 670 Ile Lys Gly Arg Phe Glu Gly Ser Arg Phe Leu Pro Tyr His Thr Arg 675 680 685 20 Asn Gln Ile Asn Gly Gly Ala Leu Asp Gly Lys Ala Pro Ile Leu Gly 690 700 Tyr Ala Glu Asp Pro Val Glu Leu Phe Phe Met His Ile Gln Gly Ser 705 710 715 720 25 Gly Arg Leu Lys Thr Pro Ser Gly Lys Tyr Ile Arg Ile Gly Tyr Ala 725 730 735 Asp Lys Asn Glu His Pro Tyr Val Ser Ile Gly Arg Tyr Met Ala Asp 740 745 75030 Lys Gly Tyr Leu Lys Leu Gly Gln Thr Ser Met Gln Gly Ile Lys Ala 755 760 765 Tyr Met Arg Gln Asn Pro Gln Arg Leu Ala Glu Val Leu Gly Gln Asn 770 775 780 35 Pro Ser Tyr Ile Phe Phe Arg Glu Leu Ala Gly Ser Ser Asn Asp Gly 785 790 800 Pro Val Gly Ala Leu Gly Thr Pro Leu Met Gly Glu Tyr Ala Gly Ala 805 810 815 40 Val Asp Arg His Tyr Ile Thr Leu Gly Ala Pro Leu Phe Val Ala Thr 820 825 830 Ala His Pro Val Thr Arg Lys Ala Leu Asn Arg Leu Ile Met Ala Gln 835 840 845 45 Asp Thr Gly Ser Ala Ile Lys Gly Ala Val Arg Val Asp Tyr Phe Trp 850 860 Gly Tyr Gly Asp Glu Ala Gly Glu Leu Ala Gly Lys Gln Lys Thr Thr 865 870 875 880 50 Gly Tyr Val Trp Gln Leu Leu Pro Asn Gly Met Lys Pro Glu Tyr Arg 885 890 895 Pro 55 <210> 96 <211> 1941

```
<212>
                                  DNA
                                  Artificial Sequence
                  <213>
                  <220>
                  <223>
                                  deltaG287NZ-953
                  <400>
                  atggctagcc ccgatgtcaa gtcggcggac acgctgtcaa aacctgccgc ccctgttgtt tctgaaaaag agacagagc aaaggaagat gcgccacagg caggttctca aggacagggc
                                                                                                                                                                  60
                                                                                                                                                                  120
                  gcgccatccg cacaaggcgg tcaagatatg gcggcggttt cggaagaaaa tacaggcaat
ggcggtgcgg cagcaacgga caaacccaaa aatgaagacg agggggcgca aaatgatatg
ccgcaaaatg ccgccgatac agatagtttg acaccgaatc acaccccggc ttcgaatatg
                                                                                                                                                                  180
                                                                                                                                                                  240
300
10
                  ccggccggaa atatggaaaa ccaagcaccg gatgccgggg aatcggagca gccggcaaac caaccggata tggcaaatac ggcggcagga atgcagggtg acgatccgtc ggcaggcgg gaaaatgccg gcaatacggc tgcccaaggt acaaatcaag ccgaaaacaa tcaaaccgcc ggttctcaaa atcctgcctc ttcaaccaat cctagcgcca cgaatagcgg tggtgatttt ggaaggacga acgtgggcaa ttctgttgtg attgacggc cgtcgcaaaa tataacgttg acccactgta aaggcgattc ttgtagtggc aataatttct tggatgaaga agtacagcta
                                                                                                                                                                  360
                                                                                                                                                                  420
                                                                                                                                                                  480
                                                                                                                                                                  540
                                                                                                                                                                  600
15
                                                                                                                                                                  660
                  aaatcagaat ttgaaaaatt aagtgatgca gacaaaataa gtaattacaa gaaagatggg
                                                                                                                                                                  720
                 aaatcagaat ttgaaaaatt aagtgatgca gacaaaataa gtaattacaa gaaagatggg aagaatgacg ggaagaatga taaatttgtc ggtttggttg ccgatagtgt gcagatgaag ggaatcaatc aatatattat cttttataaa cctaaaccca cttcatttgc gcgatttagg cgttctgcac ggtcgaggcg gtcgcttccg gccgagatgc cgctgattcc cgtcaatcag gcggatacgc tgattgtcga tggggaagcg gtcagcctga cggggcattc cggcaatatc ttcgcgcccg aagggaatta ccggtatctg acttacgggg cggaaaaatt gcccggcgga tcgtatgccc tccgtgttca aggcgaacct tcaaaaggcg aaatgctcgc gggcacggca gtgtacaacg gcgaaggtct gcatttcat acggaaaaccg gccgtccgtc cccgtccaga ggcaggttt cgcaaaagt cgatttcgc agcaaaactg tggacggcat tatcgacagc ggcgatggtt tgcatatggg tacgcaaaaa ttcaaagccg ccatcgatgg aaacggcttt aaggggactt ggacggaaaa tggcqqcqqq gatgtttccq gaaagtttta cqqccqqcc
                                                                                                                                                                 780
840
                                                                                                                                                                  900
                                                                                                                                                                  960
20
                                                                                                                                                                  1020
                                                                                                                                                                  1080
                                                                                                                                                                  1140
                                                                                                                                                                  1200
                                                                                                                                                                  1260
                 1320
25
                                                                                                                                                                  1380
                                                                                                                                                                  1440
                                                                                                                                                                 1500
1560
                                                                                                                                                                  1620
                                                                                                                                                                  1680
                                                                                                                                                                  1740
30
                                                                                                                                                                  1800
                                                                                                                                                                  1860
                                                                                                                                                                  1920
                                                                                                                                                                  1941
                  <210>
35
                                   644
                  <211>
                  <212>
                                   PRT
                                  Artificial Sequence
                  <213>
                  <220>
                  <223>
                                   deltaG287NZ-953
40
                  <400>
                  Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala
1 5 10 15
                  Ala Pro Val Val Ser Glu Lys Glu Thr Glu Ala Lys Glu Asp Ala Pro 20 25 30
45
                  Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Ala Gln Gly Gln 35 40 45
                  Asp Met Ala Ala Val Ser Glu Glu Asn Thr Gly Asn Gly Gly Ala Ala 50 55
50
                  Ala Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Ala Gln Asn Asp Met 65 70 75
                  Pro Gln Asn Ala Asp Thr Asp Ser Leu Thr Pro Asn His Thr Pro
85 90 95
55
                  Ala Ser Asn Met Pro Ala Gly Asn Met Glu Asn Gln Ala Pro Asp Ala
```

				100					105					110		
5	Gly	Glu	Ser 115	Glu	G]n	Pro	Ala	Asn 120	Gln	Pro	Asp	Met	Ala 125	Asn	Thr	Аlа
	Asp	Gly 130	Met	Gln	Gly	Asp	Asp 135	Pro	Ser	Ala	Gly	Gly 140	Glu	Asn	Ala	Gly
10	Asn 145	Thr	Ala	Ala	Gln	Gly 150	Thr	Asn	Gln	Ala	Glu 155	Asn	Asn	Gln	Thr	А]а 160
	Gly	Ser	Gln	Asn	Pro 165	Ala	Ser	Ser	Thr	Asn 170	Pro	Ser	Ala	Thr	Asn 175	Ser
15	Gly	Gly	Asp	Phe 180	Gly	Arg	Thr	Asn	Val 185	Gly	Asn	Ser	٧a٦	Val 190	Ile	Asp
15	Gly	Pro	Ser 195	Gln	Asn	Ile	Thr	Leu 200	Thr	His	Cys	Lys	Gly 205	Asp	Ser	Cys
00	Ser	Gly 210	Asn	Asn	Phe	Leu	Asp 215	Glu	Glu	Val	Gln	Leu 220	Lys	Ser	Glu	Phe
20	Glu 225	Lys	Leu	Ser	Asp	Ala 230	Asp	Lys	Ile	Ser	Asn 235	Tyr	Lys	Lys	Asp	G1y 240
	Lys	Asn	Asp	Gly	Lys 245	Asn	Asp	Lys	Phe	Va1 250	Gly	Leu	∨al	Ala	Asp 255	Ser
25	val	Gln	Met	Lys 260	Gly	Ile	Asn	Gln	Tyr 265	ıJe	Ile	Phe	Tyr	Lys 270	Pro	Lys
			275	Phe				280					285			
30		290		Glu			295					300				
	17e 305	Val	Asp	Gly	Glu	Ala 310	val	Ser	Leu	Thr	Gly 315	His	Ser	Gly	Asn	11e 320
35				Glu	325					330					335	
	_	_		G]y 340	_	-										
40	-		355	Leu		-		360		•		•	365			
	_	370	_	Glu			375					380				
45	385			Asp		390					395					400
				Leu	405		_			410					415	-
50				Phe 420	-				425					430	•	
			435	Phe				440					445			
55		450	_	Pro			455					460	_			
	Gly	Lys	Lys	Glu	Gln	Asp	Gly	Ser	Gly	Gly	Gly	Gly	Ala	Thr	Tyr	Lys

	465	470	475	480
F	Val Asp Glu Tyr His 485	Ala Asn Ala Arg Phe 490	Ala Ile Asp His Phe 495	Asn
5	Thr Ser Thr Asn Val	Gly Gly Phe Tyr Gly	Leu Thr Gly Ser Val 510	Glu
	Phe Asp Gln Ala Lys 515	Arg Asp Gly Lys Ile 520	Asp Ile Thr Ile Pro 525	Val
10	Ala Asn Leu Gln Ser 530	Gly Ser Gln His Phe 535	Thr Asp His Leu Lys 540	Ser
	Ala Asp Ile Phe Asp 545	Ala Ala Gln Tyr Pro 550	Asp Ile Arg Phe Val	ser 560
15	Thr Lys Phe Asn Phe 565	Asn Gly Lys Lys Leu 570	Val Ser Val Asp Gly 575	Asn
	Leu Thr Met His Gly 580	Lys Thr Ala Pro Val 585	Lys Leu Lys Ala Glu 590	Lys
20	Phe Asn Cys Tyr Gln 595	Ser Pro Met Ala Lys 600	Thr Glu Val Cys Gly 605	Gly
	Asp Phe Ser Thr Thr 610	Ile Asp Arg Thr Lys 615	Trp Gly Val Asp Tyr 620	Leu
25	Val Asn Val Gly Met 625	Thr Lys Ser Val Arg 630	Ile Asp Ile Gln Ile 635	Glu 640
	Ala Ala Lys Gln			
30	<210> 98 <211> 2583 <212> DNA <213> Artificial S	Sequence		
35	<220> <223> deltaG287NZ	-961		
	tctgaaaaag agacagagg gcgccatccg cacaaggc	aa gtcggcggac acgctgt gc aaaggaagat gcgccac gg tcaagatatg gcggcgg ga caaacccaaa aatgaag	agg caggttctca aggad ttt cggaagaaaa tacag	tagggc 120 ggcaat 180
40	ccgcaaaatg ccgccgata ccggccggaa atatggaaa caaccggata tggcaaata gaaaatgccg gcaatacg	ac agatagtttg acaccga aa ccaagcaccg gatgccg ac ggcggacgga atgcagg gc tgcccaaggt acaaatc tc ttcaaccaat cctagcg	atc acacccogge ttega ggg aatcggagca geegg gtg acgatecgte ggeag aag eegaaaacaa teaaa	iatatý 300 gcaaac 360 ggcggg 420 accgcc 480
45	acccactgta aaggcgattaaaatcagaat ttgaaaaataagaatgacg ggaagaatg	aa ttctgttgtg attgacg tc ttgtagtggc aataatt tt aagtgatgca gacaaaa ga taaatttgtc ggtttgg	tct tggatgaaga agtad taa gtaattacaa gaaag ttg ccgatagtgt gcaga	tagcta 660 gatggg 720 atgaag 780
50	cgttctgcac ggtcgaggg gcggatacgc tgattgtc ttcgcgcccg aagggaat tcgtatgccc tccgtgttc gtgtacaacg gcgaagtg ggcaggtttg ccgcaaaa	at cttttataaa cctaaac gg gtcgcttccg gccgaga ga tggggaagcg gtcagcc ta ccggtatctg acttacg ca aggcgaacct tcaaaag ct gcatttcat acggaaa gt cgatttcggc agcaaat	etge egetgattee egtea etga eggggeatte eggea eggg eggaaaaatt geee egeg aaatgetege gggea aeg geegteegte eeegt eetg tggaeggeat tateg	natcag 900 natatc 960 ggcgga 1020 ncggca 1080 cccaga 1140 gacagc 1200
55	aaggggačtt ggacggaa ggcgaggaag tggcgggaa ggcgtgtttg ccggcaaa	gg tacgcaaaaa ttcaaag aa tggcggcggg gatgttt aa atacagctat cgcccaa aa agagcaggat ggatccg gc cactgtggcc attgctg	ccg gaaagtttta cggco cag atgcggaaaa gggco gag gaggaggagc cacaa	ccggcc 1320 ggattc 1380 acgac 1440

atcaacggtt tcaaagctgg agagaccatc tacgacattg atgaagacgg cacaattacc

1560

```
aaaaaagacg caactgcagc cgatgttgaa gccgacgact ttaaaggtct gggtctgaaa
                                                                                                     1620
                                                                                                     1680
           aaagtcgtga ctaacctgac caaaaccgtc aatgaaaaca aacaaaacgt cgatgccaaa
           gtaaaagctg cagaatctga aatagaaaag ttaacaacca agttagcaga cactgatgcc
gctttagcag atactgatgc cgctctggat gcaaccacca acgccttgaa taaattggga
gaaaatataa cgacatttgc tgaagagact aagacaaata tcgtaaaaat tgatgaaaaa
                                                                                                     1740
                                                                                                     1800
                                                                                                     1860
           ttagaagccg tggctgatac cgtcgacaag catgccgaag cattcaacga tatcgccgat tcattggatg aaaccaacac taaggcagac gaagccgtca aaaccgccaa tgaagccaaa cagacggccg aagaaaccaa acaaaacgtc gatgccaaag taaaagctgc agaaactgca
                                                                                                     1920
                                                                                                     1980
                                                                                                     2040
           gcággcáaag ccgaagctgc cgctggcáca gctáatactg cagccgacáa ggccgaagct
                                                                                                     2100
10
           gtcgctgcaa aagttaccga catcaaagct gatatcgcta cgaacaaaga taatattgct
aaaaaagcaa acagtgccga cgtgtacacc agagaagagt ctgacagcaa atttgtcaga
                                                                                                     2160
                                                                                                     2220
                                                                                                     2280
2340
           attgatggtc tgaacgctac taccgaaaaa ttggacacac gcttggcttc tgctgaaaaa tccattgccg atcacgatac tcgcctgaac ggtttggata aaacagtgtc agacctgcgc
           aaagaaaccc gccaaggcct tgcagaacaa gccgcgctct ccggtctgtt ccaaccttac
                                                                                                     2400
           aacgtgggtc ggttcaatgt aacggctgca gtcggcggct acaaatccga atcggcagtc
                                                                                                     2460
15
           gccatcggta ccggcttccg ctttaccgaa aactttgccg ccaaagcagg cgtggcagtc
                                                                                                     2520
                                                                                                     2580
           ggcacttcgt ccggttcttc cgcagcctac catgtcggcg tcaattacga gtggtaaaag
                                                                                                     2583
                     99
858
           <210>
           <211>
           <212>
                      PRT
20
                     Artificial Sequence
           <220>
                      deltaG287NZ-961
           <223>
25
           Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala
1 10 15
           Ala Pro Val Val Ser Glu Lys Glu Thr Glu Ala Lys Glu Asp Ala Pro
20 25 30
           Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Ala Gln Gly Gln 35 40
30
           Asp Met Ala Ala Val Ser Glu Glu Asn Thr Gly Asn Gly Gly Ala Ala 50 60
           Ala Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Ala Gln Asn Asp Met 65 70 75 80
35
           Pro Gln Asn Ala Asp Thr Asp Ser Leu Thr Pro Asn His Thr Pro 85 90
           Ala Ser Asn Met Pro Ala Gly Asn Met Glu Asn Gln Ala Pro Asp Ala 100 105 110
40
           Gly Glu Ser Glu Gln Pro Ala Asn Gln Pro Asp Met Ala Asn Thr Ala 115 120 125
           Asp Gly Met Gln Gly Asp Asp Pro Ser Ala Gly Gly Glu Asn Ala Gly 130 140
45
           Asn Thr Ala Ala Gln Gly Thr Asn Gln Ala Glu Asn Asn Gln Thr Ala
145 150 160
           Gly Ser Gln Asn Pro Ala Ser Ser Thr Asn Pro Ser Ala Thr Asn Ser
165 170 175
50
           Gly Gly Asp Phe Gly Arg Thr Asn Val Gly Asn Ser Val Val Ile Asp
180 185
           Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser Cys
195 200 205
55
           Ser Gly Asn Asn Phe Leu Asp Glu Glu Val Gln Leu Lys Ser Glu Phe
```

		210					215					220				
5	Glu 225	Lys	Leu	Ser	Asp	Ala 230	Asp	Lys	Ile	Ser	Asn 235	Tyr	Lys	Lys	Asp	Gly 240
	Lys	Asn	Asp	Gly	Lys 245	Asn	Asp	Lys	Phe	va1 250	Gly	Leu	٧a٦	Ala	Asp 255	Ser
10	٧al	Gln	Met	Lys 260	Gly	Ile	Asn	Gln	Tyr 265	Ile	Ile	Phe	Tyr	Lys 270	Pro	Lys
10	Pro	Thr	Ser 275	Phe	Ala	Arg	Phe	Arg 280	Arg	Ser	Ala	Arg	Ser 285	Arg	Arg	Ser
	Leu	Pro 290	Αla	Glu	Met	Pro	Le <b>u</b> 295	Ile	Pro	۷a٦	Asn	G]n 300	Ala	Asp	Thr	Leu
15	17e 305	val	Asp	Gly	Glu	Ala 310	val	Ser	Leu	Thr	Gly 315	нis	Ser	Gly	Asn	11e 320
	Phe	Ala	Pro	Glu	Gly 325	Asn	Tyr	Arg	Tyr	Leu 330	Thr	Tyr	Gly	Αla	G]u 335	Lys
20	Leu	Pro	Gly	Gly 340	Ser	Tyr	Ala	Leu	Arg 345	Val	Gln	Gly	Glu	Pro 350	Ser	Lys
	Gly	Glu	Met 355	Leu	Ala	Gly	Thr	А]а 360	val	Tyr	Asn	Gly	Glu 365	val	Leu	His
25	Phe	ніs 370	Thr	Glu	Asn	Gly	Arg 375	Pro	Ser	Pro	Ser	Arg 380	Gly	Arg	Phe	Аlа
	Ala 385	Lys	٧a٦	Asp	Phe	Gly 390	Ser	Lys	Ser	Va1	Asp 395	Gly	Ile	Ile	Asp	Ser 400
30	Gly	Asp	Gly	Leu	ніs 405	Met	Gly	Thr	Gln	Lys 410	Phe	Lys	Ala	Ala	Ile 415	Asp
	Gly	Asn	Gly	Phe 420	Lys	Gly	Thr	Trp	Thr 425	Glu	Asn	Gly	Gly	Gly 430	Asp	٧a٦
35	Ser	Gly	Lys 435	Phe	Tyr	Gly	Pro	Ala 440	Gly	Glu	Glu	۷al	Ala 445	Gly	Lys	Tyr
	Ser	Tyr 450	Arg	Pro	Thr	Asp	Ala 455	Glu	Lys	Gly	Gly	Phe 460	Gly	val	Phe	Аlа
40	Gly 465	Lys	Lys	Glu	Gln	Asp 470	Gly	Ser	Gly	Gly	Gly 475	Gly	Ala	Thr	Asn	480
	Asp	Asp	val	Lys	Lys 485	Ala	Ala	Thr	val	Ala 490	Ile	Ala	Ala	Ala	Tyr 495	Asn
45	Asn	Gly	Gln	Glu 500	Ile	Asn	Gly	Phe	Lys 505	Ala	Gly	Glu	Thr	Ile 510	Tyr	Asp
	Ile	Asp	Glu 515	Asp	Gly	Thr	Ile	Thr 520	Lys	Lys	Asp	Ala	Thr 525	Ala	Ala	Asp
50	٧a٦	Glu 530	Ala	Asp	Asp	Phe	Lys 535	Gly	Leu	Gly	Leu	Lys 540	Lys	val	Val	Thr
	Asn 545	Leu	Thr	Lys	Thr	∨a1 550	Asn	Glu	Asn	Lys	G]n 555	Asn	val	Asp	Ala	Lys 560
55	٧a٦	Lys	Аlа	Аlа	Glu 565	Ser	Glu	Ile	Glu	Lys 570	Leu	Thr	Thr	Lys	Leu 575	Ala
	Asp	Thr	Asp	Ala	Ala	Leu	Ala	Asp	Thr	Asp	Ala	Ala	Leu	Asp	Аlа	Thr

590 580 585 Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu 595 600 605 Thr Lys Thr Asn Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val 610 620 Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp 625 630 635 10 Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala 645 650 655 Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala 660 665 670 15 Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Ala 675 680 685 Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys 690 695 700 Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn Lys Asp Asn Ile Ala 705 710 715 720 20 Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg Glu Glu Ser Asp Ser 725 730 735 Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr Thr Glu Lys Leu Asp  $740 \hspace{1.5cm} 750 \hspace{1.5cm} 750$ 25 Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala Asp His Asp Thr Arg 755 760 765 Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu Arg Lys Glu Thr Arg 770 775 780 30 Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu Phe Gln Pro Tyr 785 790 795 800 Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val Gly Gly Tyr Lys Ser 805 810 815 35 Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg Phe Thr Glu Asn Phe 820 825 830 Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser Ser Gly Ser Ser Ala 845 40 Ala Tyr His Val Gly Val Asn Tyr Glu Trp 850 855 <210> <211> 4425 45 <212> DNA Artificial Sequence <213> <220> <223> deltaG983-ORF46.1 <400> 50 atgacttctg cgcccgactt caatgcaggc ggtaccggta tcggcagcaa cagcagagca 60 acaacagcga aatcagcagc agtatcttac gccggtatca agaacgaaat gtgcaaagac agaagcatgc tctgtgccgg tcgggatgac gttgcggtta cagacaggga tgccaaaatc aatgcccccc ccccgaatct gcataccgga gactttccaa acccaaatga cgcatacaag 120 180 240 300 aatttgatca acctcaaacc tgcaattgaa gcaggctata caggacgcgg ggtagaggta ggtatcgtcg acacaggcga atccgtcggc agcatatcct ttcccgaact gtatggcaga 360 55 aaagaacacg gctataacga aaattacaaa aactatacgg cgtatatgcg gaaggaagcg 420

```
480
                       cctgaagacg gaggcggtaa agacattgaa gcttctttcg acgatgaggc cgttatagag
                       actgaagcaa agccgacgga tatccgccac gtaaaagaaa tcggacacat cgatttggtc
                                                                                                                                                                                                                                                    540
                       tcccatatta ttggcgggcg ttccgtggac ggcagacctg caggcggtat tgcgcccgat gcgacgctac acataatgaa tacgaatgat gaaaccaaga acgaaatgat ggttgcagcc
                                                                                                                                                                                                                                                    600
                                                                                                                                                                                                                                                    660
                      gcgacgctac acataatgaa tacgaatgat gaaaccaaga acgaaatgat ggttgcagcc atccgcaatg catgggtcaa gctgggcgaa cgtggcgtgc gcatcgtcaa taacagtttt ggaacaacat cgagggcagg cactgccgac cttttccaaa tagccaattc ggaggagcag taccgccaag cgttgctcga ctattccggc ggtgataaaa cagacgaggg tatccgcctg atgcaacaga gcgattacgg caacctgtcc taccacatcc gtaataaaaa catgcttttc atctttcga caggcaatga cgcacaagct cagcccaaca catatgccct attgccatt tatgaaaaag acgctcaaaa aggcattatc acagtcgcag gcgtagaccg cagtggagaa aagttcaaac gggaaatgta tggagaaccg ggtacagaac cgcttgagta tggctccaac cattgcggaa tacctgccat gtggtgcctg tcggcaccct atgaagcaag cgtccgttc accgcacaaa accgattaca aatgccgga acatccttt ccgcaccat cgtaaccggcaccg tggcggggctg tggcggaacct gcgtaccaca
 5
                                                                                                                                                                                                                                                     720
                                                                                                                                                                                                                                                    780
                                                                                                                                                                                                                                                    840
                                                                                                                                                                                                                                                     900
                                                                                                                                                                                                                                                    960
                                                                                                                                                                                                                                                    1020
10
                                                                                                                                                                                                                                                    1080
                                                                                                                                                                                                                                                    1140
                                                                                                                                                                                                                                                    1200
                       acggcggctc tgctgctgca gaaatacccg tggatgagca acgacaacct gcgtaccacg
                                                                                                                                                                                                                                                    1260
                      ttgctgacga cggctcagga catcggtgca gtcggcgtgg acagcaagtt cggctggga ctgctggatg cgggtaaggc catgaacgga cccgcgtct ttccgttcgg cgactttacc gccgatacga aaggtacatc cgatattgcc tactccttcc gtaacgacat ttcaggcacg
                                                                                                                                                                                                                                                    1320
                                                                                                                                                                                                                                                    1380
15
                                                                                                                                                                                                                                                    1440
                       ggcggcctga tcaaaaaagg cggcagccaa ctgcaactgc acggcaacaa cacctatacg
ggcaaaacca ttatcgaagg cggttcgctg gtgttgtacg gcaacaacaa atcggatatg
cgcgtcgaaa ccaaaggtgc gctgatttat aacggggcgg catccggcgg cagcctgaac
                                                                                                                                                                                                                                                    1500
                                                                                                                                                                                                                                                    1560
                                                                                                                                                                                                                                                    1620
                     cgcgtcgaaa ccaaaggtgc gctgatttat aacggggcgg catccggcgg cagcctgaac agcgacggca ttgtctatct ggcagatacc gaccaatccg gcgcaaacga aaccgtacac atcaaaggca gtctgcagct ggacggcaaa ggtacgctgt acaacggtt gggcaaactg gggggaggct atctcaacag taccggacga cgtgttccct tcctgagtgc cgccaaaatc gggcaggatt attcttctt cacaaacatc gaaaccgacg gcggcctgct ggcttccct gacagcgtcg aaaaaacagc gggcagtgaa ggcgacacgc tgtcctatta tgtccgtcgc ggcaatgcgg cacggactgc ttcggcagcg gcacattccg cgcccgccgg tctgaaacac gccgtagaac agggcggcag caatctggaa aacctgatgg tcgaactgga tgcctccgaa tactcgcaa cacccgaac ggttgaaact gcggcagcag accgcacaga tatgccgga atccgcaacagc tttccgccaa gcggcagcag tacaaccagc
                                                                                                                                                                                                                                                    1680
                                                                                                                                                                                                                                                    1740
                                                                                                                                                                                                                                                    1800
20
                                                                                                                                                                                                                                                    1860
                                                                                                                                                                                                                                                    1920
                                                                                                                                                                                                                                                    1980
                                                                                                                                                                                                                                                    2040
                                                                                                                                                                                                                                                    2100
                                                                                                                                                                                                                                                    2160
25
                       atccgcccct acggcgcaac tttccgcgca gcggcagccg tacagcatgc gaatgccgcc gacggtgtac gcatcttcaa cagtctcgcc gctaccgtct atgccgacag taccgccgcc catgccgata tgcagggacg ccgcctgaaa gccgtatcgg acgggttgga ccacaacggc
                                                                                                                                                                                                                                                    2220
                                                                                                                                                                                                                                                    2280
                                                                                                                                                                                                                                                    2340
                      acgggtctgc gcgtcatcgc gcaaacccaa caggacggtg gaacgtggga acagggcggt
gttgaaggca aaatgcgcgg cagtacccaa accgtcggca ttgccgcgaa aaccggcggaa
aatacgacag cagccgccac actgggcatg ggacgcagca catggagcga aaacagtgca
aatgcaaaaa ccgacagcat tagtctgtt gcaggcatac ggcacgatgc gggcggatatc
ggctatctca aaggcctgtt ctctacgga cgctacaaaa acagcatcag ccgcagcacc
                                                                                                                                                                                                                                                    2400
                                                                                                                                                                                                                                                    2460
                                                                                                                                                                                                                                                    2520
30
                                                                                                                                                                                                                                                    2580
                                                                                                                                                                                                                                                    2640
                       ggtgcggacg aacatgcgga aggcagcgtc aacggcacgc tgatgcagct gggcgcactg
                                                                                                                                                                                                                                                    2700
                       ggcggtgtca acgttccgtt tgccgcaacg ggagatttga cggtcgaagg cggtctgcgc
tacgacctgc tcaaacagga tgcattcgcc gaaaaaggca gtgctttggg ctggagcggc
                                                                                                                                                                                                                                                    2760
                                                                                                                                                                                                                                                    2820
                      aacagcctca ctgaaggcac gctggtcgga ctcgcgggtc tgaagctgtc gcaacccttg
agcgataaag ccgtcctgtt tgcaacggcg ggcgtggaac gcgacctgaa cggacgcgac
tacacggtaa cgggcggctt taccggcgcg actgcagcaa ccggcaagac gggggcacgc
                                                                                                                                                                                                                                                    2880
                                                                                                                                                                                                                                                    2940
35
                                                                                                                                                                                                                                                    3000
                      aatatgccgc acacccgtct ggttgccggc ctgggcgcgg atgtcgaatt cggcaacggc
tggaacggct tggcacgtta cagctacgcc ggttccaaac agtacggcaa ccacagcgga
cgagtcggcg taggctaccg gttcctcgac ggtggcggag gcactggatc ctcagatttg
                                                                                                                                                                                                                                                    3060
                                                                                                                                                                                                                                                    3120
                                                                                                                                                                                                                                                    3180
                     cgagtcggcg taggctaccg gttcctcgac ggtggcggag gcactggatc ctcagatttg gcaaacgatt ctttatccg gcaggtctc gaccgtcagc attcgaacc cgacgggaaa taccacctat tcggcagcag gggggaactt gccgagcga gcggccatat cggattggga aaaatacaaa gccatcagtt gggcaacctg atgattcaac aggcggccat taaaggaaat atcggctaca ttgtccgctt ttccgatcac gggcacgaag tccattcccc cttcgacaac catgcctcac attccgattc tgatgaagcc ggtagtcccg ttgacggatt tagcctttac cgcatccatt gggacggata cgaacaccat cccgccgacg gctatgacgg gccacagggc ggcggctatc ccgctccaa aggcgcgagg gatatataca gctacgacat aaaaggcgtt gccaaaata tccgcctcaa cctgacgac aaccgcagca ccggacaacg gcttgccgac atgccgatac agcccgatac aggcgcaggt gaacagatcg gcgacggat caaacgcgcc acccgataca gccccgagct ggacagatcg ggcaaatgccg ccgaagcctt caaacgcgcc acccgataca gccccgagct ggacagatcg ggcaaatgccg ccgaagcctt caaacgcgcc
                                                                                                                                                                                                                                                    3240
                                                                                                                                                                                                                                                    3300
40
                                                                                                                                                                                                                                                    3360
                                                                                                                                                                                                                                                    3420
                                                                                                                                                                                                                                                    3480
                                                                                                                                                                                                                                                    3540
                                                                                                                                                                                                                                                    3600
                                                                                                                                                                                                                                                    3660
45
                                                                                                                                                                                                                                                    3720
                      3780
                                                                                                                                                                                                                                                    3840
                                                                                                                                                                                                                                                    3900
                                                                                                                                                                                                                                                    3960
                      gccgcagcag ccatccgcga ttgggcagtc caaaacccca atgccgcaca aggcatagaa gccgtcagca atatctttat ggcagccatc cccatcaaag ggattggagc tgttcgggg aaatacggct tgggcggcat cacggcacat cctatcaagc ggtcgcagat ggcgcggatc gcattgccga aagggaaatc cgccgtcagc gacaattttg ccgatgcggc atacgccaaa tacccgtccc cttaccattc ccgaaatatc cgttcaaact tggagcagcg ttacggcaaa
                                                                                                                                                                                                                                                    4020
                                                                                                                                                                                                                                                    4080
50
                                                                                                                                                                                                                                                    4140
                                                                                                                                                                                                                                                    4200
                                                                                                                                                                                                                                                    4260
                      gaaaacatca cctcctcaac cgtgccgccg tcaaacggca aaaatgtcaa actggcagac caacgccacc cgaagacagg cgtaccgtt gacggtaaag ggtttccgaa ttttgagaag cacgtgaaat atgatacgct cgagcaccac caccaccacc actga
                                                                                                                                                                                                                                                    4320
                                                                                                                                                                                                                                                    4380
                                                                                                                                                                                                                                                    4425
55
```

<210> 101

	<212 <212 <213	?>	1474 PRT Arti	1 ifici	ial s	Seque	ence									
5	<220 <223		delt	taG98	33-OF	RF46	. 1									
	<400 Met 1		101 Ser	Ala	Pro 5	Asp	Phe	Asn	Ala	Gly 10	Gly	Thr	Gly	Ile	Gly 15	Ser
10	Asn	Ser	Arg	A1a 20	Thr	Thr	Аlа	Lys	Ser 25	Ala	Ala	val	Ser	Tyr 30	Ala	GТу
	Ile	Lys	Asn 35	Glu	Met	Cys	Lys	Asp 40	Arg	Ser	Met	Leu	Cys 45	Ala	Gly	Arg
15	Asp	Asp 50	٧a٦	Ala	val	Thr	Asp 55	Arg	Asp	Ala	Lys	Ile 60	Asn	Ala	Pro	Pro
	Pro 65	Asn	Leu	His	Thr	G]у 70	Asp	Phe	Pro	Asn	Pro 75	Asn	Asp	Аlа	Tyr	Lys 80
20	Asn	Leu	Ile	Asn	Leu 85	Lys	Pro	Ala	Ile	Glu 90	Ala	Gly	Tyr	Thr	G]y 95	Arg
	Gly	٧a٦	Glu	∨a1 100	Gly	Ile	val	Asp	Thr 105	GТу	Glu	Ser	Val	Gly 110	Ser	Ile
25	Ser	Phe	Pro 115	Glu	Leu	Tyr	Gly	Arg 120	Lys	Glu	His	Gly	Tyr 125	Asn	Glu	Asn
	Tyr	Lys 130	Asn	Tyr	Thr	Ala	Tyr 135	Met	Arg	Lys	Glu	Ala 140	Pro	Glu	Asp	Gly
30	Gly 145	G∃y	Lys	Asp	Ile	Glu 150	Ala	Ser	Phe	Asp	Asp 155	Glu	Ala	val	Ile	Glu 160
	Thr	Glu	Ala	Lys	Pro 165	Thr	Asp	Ile	Arg	His 170	Val	Lys	Glu	Ile	Gly 175	His
35				180					185					190	Gly	
	Pro	Ala	G]y 195	Gly	Ile	Ala	Pro	Asp 200	Ala	Thr	Leu	His	11e 205	Met	Asn	Thr
40	Asn	Asp 210	Glu	Thr	Lys	Asn	Glu 215	Met	Met	Val	Ala	Ala 220	Ile	Arg	Asn	Ala
	225		-			230	_			_	235				Ser	240
45	Gly	Thr	Thr	Ser	Arg 245	Ala	Gly	Thr	Ala	Asp 250	Leu	Phe	Gln	Ile	Ala 255	Asn
	Ser	Glu	Glu	G]n 260	Tyr	Arg	Gln	Ala	Leu 265	Leu	Asp	Tyr	Ser	Gly 270	Gly	Asp
50	Lys	Thr	Asp 275	Glu	Gly	Ile	Arg	Leu 280	Met	Gln	Gln	Ser	Asp 285	Tyr	Gly	Asn
		290	-			_	295					300			Ser	
55	305					310					315				Pro	320
	Tyr	Glu	Lys	Asp	Ala	Gln	Lys	Gly	Ile	Ile	Thr	۷a٦	Ala	Gly	val	Asp

					325					330					335	
5	Arg	Ser	Gly	G1u 340	Lys	Phe	Lys	Arg	Glu 345	Met	Tyr	Gly	Glu	Pro 350	Gly	Thr
	Glu	Pro	Leu 355	Glu	Tyr	Gly	ser	Asn 360	ніѕ	Cys	Gly	ıle	Thr 365	Ala	Met	Тгр
10	Cys	Leu 370	Ser	Ala	Pro	Tyr	G] u 375	Ala	Ser	va1	Arg	Phe 380	Thr	Arg	Thr	Asn
	Pro 385	Ile	Gln	Ile	Ala	Gly 390	Thr	Ser	Phe	Ser	Ala 395	Pro	Ile	val	Thr	G⊺y 400
15					405			Lys	-	410	·				415	
		_		420				Thr	425					430		
20			435			_		Gly 440					445			
		450					455	Phe	_			460				_
25	465					470	-	Ser		_	475	·			-	480
25	•	-			485	•	-	Gly		490					495	
00				500				Ile	505					510		
30			515					Met 520	_				525	-		
		530					535	Gly				540				
35	545	-	_			550		Gln		_	555					560
					565			Asp Asp	-	570	_				575	
40				580		-		Lys	585					590		-
		-	595					600	-			•	605			
45	•	61Ō	_				615	Ser			-	620	-		-	-
	625					630		Thr			635					640
50					645			Gly		650					655	
				660				Ala	665					670		
55			675		-			ніs 680		_			685			
	Leu	GIU	ASN	Leu	мет	val	GIU	Leu	ASP	АIA	ser	GIU	ser	ser	Ala	ınr

690 695 700 Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr Asp Met Pro Gly 705 715 720 Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Ala Val Gln His 725 730 735 Ala Asn Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala Ala Thr 740 745 750 10 Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly Arg Arg 755 760 765 Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly Leu Arg 770 780 15 Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln Gly Gly 785 790 795 800 Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala 805 810 815 20 Lys Thr Gly Glu Asn Thr Thr Ala Ala Ala Thr Leu Gly Met Gly Arg 820 825 830 Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser 835 840 845 Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys 850 860 25 Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr 865 870 875 Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln 885 890 895 30 Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp 900 905 910 Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys Gln Asp Ala 915 920 925 35 Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser Leu Thr 930 935 940 Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln Pro Leu 945 950 955 960 40 Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val Glu Arg Asp Leu 965 970 975 Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly Ala Thr Ala 980 985 990 45 Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His Thr Arg Leu Val 995 1000 1005 Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly Trp Asn Gly Leu 1010 1020 50 Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn His Ser Gly 1025 1030 1035 1040 Arg Val Gly Val Gly Tyr Arg Phe Leu Asp Gly Gly Gly Gly Thr Gly 1045 1050 1055 55 Ser Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg

				1060	)				1069	5				1070	)	
5	Gln	His	Phe 1075		Pro	Asp	Gly	Lys 1080		His	Leu	Phe	Gly 1085		Arg	Gly
J	Glu	Leu 1090		Glu	Arg	Ser	Gly 1095		Ile	Gly	Leu	Gly 1100		Ile	Gln	Ser
10	His 1105		Leu	Gly	Asn	Leu 1110		Ile	Gln	Gln	Ala 1115		Ile	Lys	Gly	Asn 1120
10	Ile	Gly	Tyr	Ile	va7 1125		Phe	Ser	Asp	ніs 1130		His	Glu	Val	ніs 1139	
	Pro	Phe	Asp	Asn 114(	нis )	Ala	Ser	His	Ser 1145		Ser	Asp	Glu	Ala 1150		Ser
15	Pro	val	Asp 1155		Phe	ser	Leu	Tyr 1160		Ile	His	Trp	Asp 1165		Tyr	Glu
	His	His 1170		Ala	Asp	Glу	Tyr 1175		Gly	Pro	Gln	Gly 1180		Gly	Tyr	Pro
20	Ala 1185		Lys	Gly	Ala	Arg 1190		Ile	Tyr	Ser	Tyr 119		Ile	Lys	Gly	∨a1 1200
	Ala	Gln	Asn	Ile	Arg 1205		Asn	Leu	Thr	Asp 121(		Arg	Ser	Thr	Gly 121	
25	Arg	Leu	Ala	Asp 1220	Arg )	Phe	His	Asn	Ala 1229		Ser	Met	Leu	Thr 123(		Gly
	Val	Gly	Asp 1235		Phe	Lys	Arg	Ala 1240		Arg	Tyr	Ser	Pro 1245		Leu	Asp
30	Arg	Ser 1250	Gly )	Asn	Ala	Ala	Glu 1255		Phe	Asn	Gly	Thr 1260		Asp	Ile	val
	1265	5			Gly	1270	)	-			1275	5		-		1280
35					Ser 1285	5				1290	)				1299	5
	•			1300				-	1305	5				1310	)	
40			1315	5	Leu	•		1320	)				1325	5	•	
		1330	)		Pro		1335	)				1340	)			
45	11e 1345		Met	Ala	Ala	1]e 1350		Ile	Lys	Gly	1355		ΑΊа	val	Arg	Gly 1360
	Lys	Tyr	Gly	Leu	Gly 1365		Ile	Thr	Ala	ніs 1370		Ile	Lys	Arg	Ser 1379	
50	Met	Gly	Ala	11e 1380	Ala )	Leu	Pro	Lys	Gly 1385		Ser	Ala	۷a٦	Ser 1390		Asn
			1395	5	Ala			1400	)				1405	5		
55		1410	)		Asn		1415	5				1420	)			
	Ser	Ser	Thr	٧a٦	Pro	Pro	Ser	Asn	Gly	Lys	Asn	٧a٦	Lys	Leu	Ala	Asp

```
1425
                                                                       1430
                                                                                                                        1435
                                                                                                                                                                          1440
                     Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro
1445 1450 1455
                     Asn Phe Glu Lys His Val Lys Tyr Asp Thr Leu Glu His His His 1460 1465 1470
                     His His
10
                     <210>
                                         102
                                         3939
                     <211>
                      <212>
                                         DNA
                                         Artificial Sequence
                      <213>
15
                      <220>
                      <223>
                                         deltaG983-741
                      <400>
                     atgacttctg cgcccgactt caatgcaggc ggtaccggta tcggcagcaa cagcagagca acaacagcga aatcagcagc agtatcttac gccggtatca agaacgaaat gtgcaaagac
                                                                                                                                                                                              60
                                                                                                                                                                                              120
                     agaagcatgc tctgtgccgg tcgggatgac gttgcggtta cagacaggga tgccaaaatc
                                                                                                                                                                                              180
20
                     aatgccccc cccgaatct gcataccgga gactttccaa acccaaatga cgcatacaag aatttgatca acctcaaacc tgcaattgaa gcaggctata caggacgcgg ggtagaggta ggtatcgtcg acacaggcga atccgtcggc agcatatcct ttcccgaact gtatggcaga
                                                                                                                                                                                              240
                                                                                                                                                                                              300
                                                                                                                                                                                              360
                     aaagaacacg gctataacga aaattacaaa aactatacgg cgtatatgcg gaaggaagcg cctgaagacg gaggcggtaa agacattgaa gcttctttcg acgatgaggc cgttatagag actgaagcaa agccgacgga tatccgccac gtaaaagaaa tcggacacat cgatttggtc tcccatatta ttggcgggcg ttccgtggac ggcagacctg caggcggtat tgcgcccgat gcgacgctac acataatgaa tacgaatgat gaaaccaaga acgaaatgat ggttgcagcc
                                                                                                                                                                                              420
480
                                                                                                                                                                                              540
25
                                                                                                                                                                                              600
                                                                                                                                                                                              660
                     atccgcaatg catgggtcaa gctgggcgaa cgtggcgtgc gcatcgtcaa taacagttt
ggaacaacat cgagggcagg cactgccgac cttttccaaa tagccaattc ggaggagcag
taccgccaag cgttgctcga ctattccggc ggtgataaaa cagacgaggg tatccgcctg
atgcaacaga gcgattacgg caacctgtcc taccacatcc gtaataaaaa catgcttttc
                                                                                                                                                                                              720
                                                                                                                                                                                             780
840
                                                                                                                                                                                              900
                     atcttttcga caggcaatga cgcacaagct cagcccaaca catatgccct attgccattt tatgaaaaaag acgctcaaaa aggcattatc acagtcgcag gcgtagaccg cagtggagaa aagttcaaac gggaaatgta tggagaaccg ggtacagaac cgcttgagta tggctccaac cattgcggaa ttactgccat gtggtgcctg tcggcaccct atgaagcaag cgtccgttc acccgtacaa acccgatca aattgccgga acattccttt ccgcacccat cgtaaccggc
30
                                                                                                                                                                                              960
                                                                                                                                                                                              1020
                                                                                                                                                                                              1080
                                                                                                                                                                                              1140
                                                                                                                                                                                              1200
                     acggcggctc tgctgctgca gaaatacccg tggatgagca acgacaacct gcgtaccacg
                                                                                                                                                                                              1260
                     ttgctgacga cggctcagga catcggtgca gtcggcgtgg acagcaagtt cggctgggga ctgctggatg cgggtaaggc catgaacgga cccgcgtct ttccgttcgg cgactttacc gccgatacga aaggtacatc cgatattgcc tactccttcc gtaacgacat ttcaggcacg
                                                                                                                                                                                              1320
1380
35
                                                                                                                                                                                              1440
                    gccgatacga aaggtacatc cgatattgcc tactccttcc gtaacgacat ttcaggcacg ggcggcctga tcaaaaaagg cggcagccaa ctgcaactgc acggcaacaa cacctatacg ggcaaaacca ttatcgaagg cggttcgctg gtgttgtacg gcaacaacaa atcggatatg cggttcgaaa ccaaaggtgc gctgatttat aacggggcgg catccggcgg cagcctgaac agcgacggca ttgtctatct ggcagatacc gaccaatccg gcgcaaacga aaccgtacac atcaaaggca gtctgcagct ggacggcaaa ggtacgctgt acacacgttt gggcaaactg ctgaaagtgg acggtacggc gattatcggc ggcaagctgt acatgtcggc acgcggcaag ggggcaggct atctcaacag taccggacga cgtgttccct tcctgagtgc cgccaaaatc gggcaggatt atctttctt cacaaacacac gaaaccgacg gcggcctgct ggcttccct
                                                                                                                                                                                              1500
                                                                                                                                                                                              1560
                                                                                                                                                                                              1620
                                                                                                                                                                                              1680
40
                                                                                                                                                                                              1740
                                                                                                                                                                                              1800
                                                                                                                                                                                              1860
                                                                                                                                                                                             1920
                     gacagcgtcg aaaaaacagc gggcagtgaa ggcgacacgc tgtcctatta tgtccgtcgc ggcaatgcgg cacggactgc ttcggcagcg gcacattccg cgcccgccgg tctgaaacac gccgtagaac agggcggcag caatctggaa aacctgatgg tcgaactgga tgcctccgaa
                                                                                                                                                                                              1980
                                                                                                                                                                                              2040
45
                                                                                                                                                                                              2100
                     tcatccgcaa cacccgagac ggttgaaact gcggcagccg accgcacaga tatgccggc atccgccct acggcgcaac tttccgcgca gcggcagccg tacagcatgc gaatgccgc gacggtgtac gcatcttcaa cagtctcgcc gctaccgtct atgccgacag taccgccgcc
                                                                                                                                                                                              2160
                                                                                                                                                                                              2220
2280
                      catóccoata tocagogaco coocctoaaa occotatego acogottoga ccacaacogo
                                                                                                                                                                                              2340
                     acgggtctgc gcgtcatcgc gcaaacccaa caggacggtg gaacgtggga acagggcggt
gttgaaggca aaatgcgcgg cagtacccaa accgtcggca ttgccgcgaa aaccggcgaa
                                                                                                                                                                                              2400
                                                                                                                                                                                              2460
50
                     aatacgacag cagccgccac actgggcatg ggacgcagca catggagcga aaacagtgca aatgcaaaaa ccgacagcat tagtctgttt gcaggcatac ggcacgatgc gggcgatatc ggctatctca aaggcctgtt ctcctacgga cgctacaaaa acagcatcag ccgcagcacc ggtgcggacg aacatgcgga aggcagcgtc aacggcacgc tgatgcagct gggcgcactg
                                                                                                                                                                                              2520
                                                                                                                                                                                              2580
                                                                                                                                                                                              2640
                                                                                                                                                                                              2700
                     ggcggtgtca acgttccgtt tgccgcaacg ggagatttga cggtcgaagg cggtctgcgc
tacgacctgc tcaaacagga tgcattcgcc gaaaaaggca gtgctttggg ctggagcggc
                                                                                                                                                                                              2760
                                                                                                                                                                                              2820
55
                     aacageetea etgaaggeae getggtegga etegegggte tgaagetgte geaaceettg
                                                                                                                                                                                              2880
                     agcgataaag ccgtcctgtt tgcaacggcg ggcgtggaac gcgacctgaa cggacgcgac
                                                                                                                                                                                              2940
```

3000 3060

3120

3180 3240

3300 3360

3420 3480

3540

3600 3660

3720

3780 3840

3900 3939

```
tacacggtaa cgggcggctt taccggcgcg actgcagcaa ccggcaagac gggggcacgc
aatatgccgc acacccgtct ggttgccggc ctgggcgcgg atgtcgaatt cggcaacggc
          tggaacggčt tggcacgtta čagctacgčc ggťťčcaaač agťacggcaa ccacagcgga
          cgagtcggcg taggctaccg gticctcgag ggatccggag ggggtggtgt cgccgccgacatcggtgcgg ggcttgccga tgcactaacc gcaccgctcg accataaaga caaaggtttg
          cagicitiga cociggatca giccgicago aaaaacgaga aacigaagci ggcggcacaa
ggigcggaaa aaacitaigg aaacggigac agccicaata cgggcaaati gaagaacgac
          aaggtcagcc gtttcgactt tatccgccaa atcgaagtgg acgggcagct cattaccttg
gagagtggag agttccaagt atacaaacaa agccattccg ccttaaccgc ctttcagacc
          gagcaaatac aagattegga geatteeggg aagatggttg egaaaegeca gtteagaate
          ggcgacatag cgggcgaaca tacatctttt gacaagcttc ccgaaggcgg cagggcgaca
tatcgcggga cggcgttcgg ttcagacgat gccggcggaa aactgaccta caccatagat
10
          ttcgčcgčča agčagggaãa cggcãaaatc gaacattīga aatcgccaga actcaatgtc
          gacctggccg ccgccgatat caagccggat ggaaaacgcc atgccgtcat cagcggttcc
gtcctttaca accaagccga gaaaggcagt tactccctcg gtatctttgg cggaaaagcc
          caggaagttg ccggcagcgc ggaagtgaaa accgtaaacg gcatacgcca tatcggcctt
          gccgccaagc aactcgagca ccaccaccac caccactga
          <210>
                     103
                    1312
          <211>
          <212>
                    PRT
                    Artificial Sequence
          <213>
20
          <220>
<223>
                    deltaG983-741
          <400>
         Met Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser 1 5 10
25
         Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly
20 25 30
          Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg
30
         Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro 50 60
          Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys 70 75 80
35
         Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg
85 90 95
         Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile 100 \hspace{1.5cm} 105 \hspace{1.5cm} 10
              Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn
115 120 125
         Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly 130 140
         Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu
145 150 160
          Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His
165 170 175
50
         Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg 180 185 190
         Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr
195 200 205
         Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala 210 220
```

Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe 225 230 235 Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn 245 250 255 Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp  $260 \hspace{1.5cm} 265 \hspace{1.5cm} 270 \hspace{1.5cm}$ Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn 275 280 285 10 Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr 290 295 300 Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe 305 310 315 15 Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp 325 330 335 Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr  $340 \hspace{1.5cm} 345 \hspace{1.5cm} 350$ 20 Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp 355 360 365 Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn 370 375 380 25 Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly 385 390 400 Thr Ala Ala Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn Asp Asn 405 410 41530 Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly 420 425 430 Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys Ala Met 435 440 35 Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp Thr Lys 450 460 Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser Gly Thr 465 470 475 480 40 Gly Gly Leu Ile Lys Lys Gly Gly Ser Gln Leu Gln Leu His Gly Asn 485 490 495 Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly Ser Leu Val Leu 500 505 510 45 Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly Ala Leu 515 520 525 Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp Gly Ile 530 540 50 Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His 545 550 555 560 Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg 565 570 575 Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys 580 585 590 55

Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr 595 600 605Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr 610 620 Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala Ser Leu 625 630 635 640 Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu Ser Tyr 645 650 655 10 Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala Ala His 660 665 670 Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln Gly Gly Ser Asn 675 680 685 15 Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu Ser Ser Ala Thr 690 695 700 Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr Asp Met Pro Gly 705 710 720 20 Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Ala Val Gln His 725 730 735 Ala Asn Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala Ala Thr 740 745 750 25 Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly Arg Arg 755 760 765 Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly Leu Arg 770 775 780 30 Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln Gly 785 790 795 Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala 805 810 815 35 Lys Thr Gly Glu Asn Thr Thr Ala Ala Ala Thr Leu Gly Met Gly Arg 820 825 830 Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser 835 840 845 40 Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys 850 860 Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr 865 870 875 45 Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln 885 890 Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp 900 905 910 Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys Gln Asp Ala 915 920 925 50 Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser Leu Thr 930 940 Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln Pro Leu 945 950 955 960 55

	Ser Asp	Lys Ala	Val Le 965	eu Phe	Ala	Thr	Ala 970	Gly	∨a1	Glu	Arg	Asp 975	Leu
5	Asn Gly	Arg Asp 980	Tyr T	hr Val		G]y 985	Gly	Phe	Thr	Glу	Ala 990	Thr	Ala
	Ala Thr	Gly Lys 995	Thr G	ly Ala	Arg 1000		Met	Pro	His	Thr 1005		Leu	val
10	Ala Gly 1010	Leu Gly O	Ala A	sp Val 1015		Phe	Gly	Asn	Gly 1020		Asn	Gly	Leu
	Ala Arg 1025	Tyr Ser		la Gly 030	Ser	Lys	Gln	Tyr 1035		Asn	His	Ser	Gly 1040
15	Arg Val	Gly Val	Gly T	yr Arg	Phe	Leu	Glu 1050		Ser	Gly	Gly	Gly 1055	
	Val Ala	Ala Asp 106		ly Ala		Leu 1065		Asp	Аlа	Leu	Thr 1070		Pro
20	Leu Asp	His Lys 1075	Asp L	ys Gly	Leu 1080		Ser	Leu	Thr	Leu 1085		Gln	Ser
	val Arg 1090	Lys Asn O	Glu L	ys Leu 1095		Leu	Ala	Ala	G]n 1100		Ala	Glu	Lys
25	Thr Tyr 1105	Gly Asn		sp Ser 110	Leu	Asn	Thr	Gly 1115		Leu	Lys	Asn	Asp 1120
	Lys Val	Ser Arg	Phe A: 1125	sp Phe	Ile	Arg	Gln 1130		Glu	٧a٦	Asp	Gly 1135	
00	Leu Ile	Thr Leu 114		er Gly		Phe 1145		val	Tyr	Lys	G]n 1150		His
30	Ser Ala	Leu Thr 1155	Ala P	he Gln	Thr 1160		Gln	Ile	Gln	Asp 1165		Glu	His
	Ser Gly 1170	Lys Met O	Val A	la Lys 1175		Gln	Phe	Arg	Ile 1180		Asp	Ile	Ala
35	Gly Glu 1185	His Thr		he Asp 190	Lys	Leu	Pro	Glu 1195		Gly	Arg	Ala	Thr 1200
	Tyr Arg	Gly Thr	Ala Pl 1205	he Gly	Ser	Asp	Asp 1210		Gly	Glу	Lys	Leu 1215	
40	Tyr Thr	Ile Asp 1220		la Ala		Gln 1225		Asn	Gly	Lys	11e 1230		His
	Leu Lys	Ser Pro 1235	Glu L	eu Asn	val 1240		Leu	Ala	Ala	Ala 1249		Ile	Lys
45	Pro Asp 1250	Gly Lys )	Arg H	is Ala 1255		Ile	Ser	Gly	Ser 1260		Leu	Tyr	Asn
	Gln Ala 1265	Glu Lys		er Tyr 270	Ser	Leu	Gly	11e 1275		Glу	Gly	Lys	Ala 1280
50	Gln Glu	Val Ala	Gly Se 1285	er Ala	Glu	val	Lys 1290		٧a٦	Asn	Gly	11e 1295	
	His Ile	Gly Leu 130		la Lys		Leu 1305		ніѕ	His	His	His 1310		His
55	<210> <211>	104 4344											

```
<212>
                                       DNA
                    <213>
                                       Artificial Sequence
                   <220>
                   <223>
                                       deltaG983-961
                   <400>
                                       104
                   atgacttctg cgcccgactt caatgcaggc ggtaccggta tcggcagcaa cagcagagca
                   acăacagcgă aatcagcagc agtatcttac gccggtatca agaacgaaat gtgcaaagac
                                                                                                                                                                                               120
                                                                                                                                                                                               180
                   agaagcatgc tctgtgccgg tcgggatgac gttgcggtta cagacaggga tgccaaaatc
                   aatgccccc cccgaatct gcataccgga gactttccaa acccaaatga cgcatacaag aatttgatca acctcaaacc tgcaattgaa gcaggctata caggacgcgg ggtagaggta ggtatcgtcg acacaggcga atccgtcggc agcatatcct ttcccgaact gtatggcaga
10
                                                                                                                                                                                               240
                                                                                                                                                                                               300
                                                                                                                                                                                               360
                   aaagaacacg gctataacga aaattacaaa aactatacgg cgtatatgcg gaaggaagcg
cctgaagacg gaggcggtaa agacattgaa gcttctttcg acgatgaggc cgttatagag
                                                                                                                                                                                               420
                                                                                                                                                                                               480
                   actgaagcaa agccgacgga tatccgccac gtaaaagaaa tcggacacat cgatttggtc
                                                                                                                                                                                               540
                   tcccatatta ttggcgggcg ttccgtggac ggcagacctg caggcggtat tgcgcccgat gcgacgctac acataatgaa tacgaatgat gaaaccaaga acgaaatgat ggttgcagcc
                                                                                                                                                                                               600
                                                                                                                                                                                               660
                   atccgcaatg catgggtcaa gctgggcgaa cgtggcgtgc gcatcgtcaa taacagtttt ggaacaacat cgagggcagg cactgccgac cttttccaaa tagccaattc ggaggagcag taccgccaag cgttgctcga ctattccggc ggtgataaaa cagacgaggg tatccgcctg atgcaacaga gcgattacgg caacctgtcc taccacatcc gtaataaaaa catgcttttc
                                                                                                                                                                                               720
                                                                                                                                                                                               780
                                                                                                                                                                                               840
                                                                                                                                                                                               900
                  atcttttcga caggcaatga cgcacaagct cagcccaaca catatgccct attgccattt
tatgaaaaag acgctcaaaa aggcattatc acagtcgcag gcgtagaccg cagtggagaa
aagttcaaac gggaaatgta tggagaaccg ggtacagaac cgcttgagta tggctccaac
cattgcggaa ttactgccat gtggtgcctg tcggcaccct atgaagcaag cgtccgtttc
acccgtacaa acccgattca aattgccgga acatcctttt ccgcacccat cgtaaccggc
                                                                                                                                                                                               960
20
                                                                                                                                                                                               1020
                                                                                                                                                                                               1080
                                                                                                                                                                                              1140
                                                                                                                                                                                              1200
                   acggcggctc tgctgctgca gaaatacccg tggatgagca acgacaacct gcgtaccacg
                                                                                                                                                                                              1260
                   ttgctgacga cggctcagga catcggtgca gtcggcgtgg acagcaagtt cggctgggga ctgctggatg cgggtaaggc catgaacgga cccgcgtcct ttccgttcgg cgactttacc
                                                                                                                                                                                               1320
25
                                                                                                                                                                                              1380
                   gccgatacga aaggtacate cgatattgce tacteettee gtaacgacat ttcaggcacg
                                                                                                                                                                                              1440
                   ggcggcctga tcaaaaaagg cggcagccaa ctgcaactgc acggcaacaa cacctatacg
ggcaaaacca ttatcgaagg cggttcgctg gtgttgtacg gcaacaacaa atcggatatg
cgcgtcgaaa ccaaaggtgc gctgatttat aacggggcgg catccggcgg cagcctgaac
                                                                                                                                                                                              1500
                                                                                                                                                                                              1560
                                                                                                                                                                                               1620
                  cgcgtcgaaa ccaaaggtgc gctgatttat aacggggcgg catccggcgg cagcctgaac agcgacggca ttgtctatct ggcagatacc gaccaatccg gcgcaaacga aaccgtacac atcaaaggca gtctgcagct ggacggcaaa ggtacgctgt acacacgtt gggcaaactg ggggcaagtgg acggtacgcg gattatcggc ggcaagctgt acatgtcggc acgcggcaag ggggcaggct atctcaacag taccggacga cgtgttccct tcctgagtgc cgccaaaatc gggcaggatt attcttctt cacaaacatc gaaaccgacg gcggcctgct ggcttccctc gacagcgtcg aaaaaacagc gggcagtgaa ggcgacacgc tgtcctatta tgtccgtcgc ggcaatgcgg cacggactgc ttcggcagcg gcacattccg cgcccgccgg tctgaaacac gccgtagaac agggcggcag caatctggaa aacctgatgg tcgaactgga tgcctccgaa tacccgcaac acccgcaacac gtttccgaacac gcggcagcag accgcaacac taccgcaacac acccgcaacac tttccgcaacac gcggcagcag accatac gaaaccatac gaaaccacac
                                                                                                                                                                                              1680
                                                                                                                                                                                               1740
30
                                                                                                                                                                                               1800
                                                                                                                                                                                              1860
                                                                                                                                                                                               1920
                                                                                                                                                                                               1980
                                                                                                                                                                                               2040
                                                                                                                                                                                               2100
35
                                                                                                                                                                                               2160
                   atccgccct acggcgcaac titccgcgca gcggcagccg tacagcatgc gaatgccgcc gacggtgtac gcatcttcaa cagtctcgcc gctaccgtct atgccgacag taccgccgcc
                                                                                                                                                                                              2220
2280
                   čatýčcýata tycagygacy ccýcctyáaa ýccytatcyy acygyttygá ccacáacyyc
                                                                                                                                                                                               2340
                   acgggtctgc gcgtcatcgc gcaaacccaa caggacggtg gaacgtggga acagggcggt
gttgaaggca aaatgcgcgg cagtacccaa accgtcggca ttgccgcgaa aaccggcgga
aatacgacag cagccgccac actggggatg ggacgacac catggagga aaacagtgca
                                                                                                                                                                                               2400
                                                                                                                                                                                               2460
40
                                                                                                                                                                                               2520
                   aatgcaaaaa ccgacagcat tagtctgttt gcaggcatac ggcacgatgc gggcgatatc ggctatctca aaggcctgtt ctcctacgga cgctacaaaa acagcatcag ccgcagcacc
                                                                                                                                                                                               2580
                                                                                                                                                                                               2640
                   ggtgcggacg aacatgcgga aggcagcgtc aacggcacgc tgatgcagct gggcgcactg
ggcggtgtca acgttccgtt tgccgcaacg ggagatttga cggtcgaagg cggtctgcgc
tacgacctgc tcaaacagga tgcattcgcc gaaaaaggca gtgctttggg ctggagcggc
aacagcctca ctgaaggcac gctggtcgga ctcgcgggtc tgaagctgtc gcaacccttg
                                                                                                                                                                                               2700
                                                                                                                                                                                               2760
                                                                                                                                                                                               2820
                                                                                                                                                                                               2880
45
                   agcgataaag ccgtcctgtt tgcaacggcg ggcgtggaac gcgacctgaa cggacgcgac tacacggtaa cgggcggctt taccggcgcg actgcagcaa ccggcaagac gggggcacgc aatatgccgc acacccgtct ggttgccggc ctgggcgcgg atgtcgaatt cggcaacggc tggaacggct tggcacgtta cagctacgcc ggttccaaac agtacggcaa ccacagcgga
                                                                                                                                                                                               2940
                                                                                                                                                                                               3000
                                                                                                                                                                                               3060
                                                                                                                                                                                               3120
                  cgagtcggcg taggctaccg gttcctcgag ggtggcggag gcactggatc cgccacaaac gacgacgatg ttaaaaaagc tgccactgtg gccattgctg ctgcctacaa caatggccaa gaaatcaacg gttcaaagc tggagagacc atctacgaca ttgatgaaga cggcacaatt accaaaaaag acgcaactgc agccgatgtt gaagccgacg actttaaagg tctgggtctg aaaaaaagtcg tgactaacct gaccaaaaacc gtcaatgaaa acaaacaaaa cgtcgatgc
                                                                                                                                                                                               3180
                                                                                                                                                                                               3240
50
                                                                                                                                                                                               3300
                                                                                                                                                                                               3360
                                                                                                                                                                                               3420
                   aaagtaaaag ctgcagaatc tgaaatagaa aagttaacaa ccaagttagc agacactgat
                                                                                                                                                                                               3480
                   gccgctttag cagatactga tgccgctctg gatgcaacca ccaacgcctt gaataaattg ggagaaaata taacgacatt tgctgaagag actaagacaa atatcgtaaa aattagatga aaattagaag ccgtggctga taccgtcgac aagcatgccg aagcattcaa cgatatcgcc gattcattgg atgaaaccaa cactaaggca gacgaagccg tcaaaaccgc caatgaagcc
                                                                                                                                                                                               3540
                                                                                                                                                                                               3600
55
                                                                                                                                                                                               3660
                                                                                                                                                                                               3720
```

3780

aaacagacgg ccgaagaaac caaacaaaac gtcgatgcca aagtaaaagc tgcagaaact

```
gcagcaggca aagccgaagc tgccgctggc acagctaata ctgcagccga caaggccgaa
                                                                                                3840
                                                                                                3900
          gctgtcgctg caaaagttac cgacatcaaa gctgatatcg ctacgaacaa agataatatt
          gctaaaaaag caaacagtgc cgacgtgtac accagagaag agtctgacag caaatttgtc
                                                                                                3960
          agaattgatǧ gtctgaacǧc tactaccgaa aaattgǧaca cacgcttggč ttctgctǧaa
                                                                                                4020
          aaatccattg ccgatcacga tactcgcctg aacggtttgg ataaaacagt gtcagacctg cgcaaagaaa cccgccaagg ccttgcagaa caagccgcgc tctccggtct gttccaacct
                                                                                                4080
                                                                                                4140
          tacaacgtgg gtcggttcaa tgtaacggct gcagtcggcg gctacaaatc cgaatcggca
gtcgccatcg gtaccggctt ccgctttacc gaaaactttg ccgccaaagc aggcgtggca
gtcggcactt cgtccggttc ttccgcagcc taccatgtcg gcgtcaatta cgagtggctc
                                                                                                4200
                                                                                                4260
                                                                                                4320
10
          gagcaccacc accaccacca ctga
                                                                                                4344
           <210>
                    105
          <211>
                    1447
          <212>
                    PRT
                    Artificial Sequence
           <213>
15
          <220>
                    deltaG983-961
          <223>
          <400>
          Met Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser
1 10 15
20
          Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly
20 25 30
          Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg 35 40 45
25
          Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro 50 55 60
          Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys 70 75 80
30
          Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg
85 90 95
          Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile
100 105 110
35
          Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn
115 120 125
          Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly 130 135
40
          Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu
145 150 155
          Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His
165 170 175
          Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg 180 185 190
          Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr
195 200 205
          Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala 210 220
          Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe 225 230 235
          Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn 250 255
```

Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp 260 265 270 Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn 275 280 285 Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr 290 295 300 Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe 305 310 315 320 10 Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp 325 330 335 Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr  $340 \hspace{1.5cm} 345 \hspace{1.5cm} 350$ 15 Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp  $355 \hspace{1.5cm} 360 \hspace{1.5cm} 365$ Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn 370 380 20 Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly 385 390 395 400 Thr Ala Ala Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn Asp Asn 405 410 415 25 Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly 420 425 430 30 Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp Thr Lys 450 455 460 Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser Gly Thr 465 470 480 35 Gly Gly Leu Ile Lys Lys Gly Gly Ser Gln Leu Gln Leu His Gly Asn  $485 \hspace{1.5cm} 490 \hspace{1.5cm} 495$ Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly Ser Leu Val Leu 500 505 510 40 Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly Ala Leu 515 520 525 Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp Gly Ile 530 540 45 Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His 545 550 555 560 Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg 565 570 575 50 Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys 580 585 Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr 595 600 Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr 610 620

	Ser 625	Phe	Phe	Thr	Asn	Ile 630	Glu	Thr	Asp	Gly	G]y 635	Leu	Leu	Аlа	Ser	Leu 640
5	Asp	Ser	Val	Glu	Lys 645	Thr	Ala	Gly	Ser	Glu 650	Gly	Asp	Thr	Leu	Ser 655	Tyr
	Tyr	٧a٦	Arg	Arg 660	Gly	Asn	Ala	Ala	Arg 665	Thr	Ala	Ser	Ala	А]а 670	Ala	His
10	Ser	Ala	Pro 675	Ala	Gly	Leu	Lys	ніs 680	Ala	Val	Glu	Gln	Gly 685	Gly	Ser	Asn
	Leu	Glu 690	Asn	Leu	Met	۷al	Glu 695	Leu	Asp	Ala	Ser	Glu 700	Ser	Ser	Ala	Thr
15	Pro 705	Glu	Thr	val	Glu	Thr 710	Ala	Ala	Ala	Asp	Arg 715	Thr	Asp	Met	Pro	G]y 720
	Ile	Arg	Pro	Tyr	G∏y 725	Ala	Thr	Phe	Arg	Ala 730	Ala	Ala	Ala	val	G]n 735	His
20	Ala	Asn	Ala	Ala 740	Asp	Gly	Val	Arg	11e 745	Phe	Asn	Ser	Leu	Ala 750	Ala	Thr
	val	Tyr	Ala 755	Asp	Ser	Thr	Ala	Ala 760	ніѕ	Ala	Asp	Met	G1n 765	Gly	Arg	Arg
25	Leu	Lys 770	Ala	val	Ser	Asp	Gly 775	Leu	Asp	His	Asn	Gly 780	Thr	Gly	Leu	Arg
	Val 785	Ile	Ala	Gln	Thr	G]n 790	Gln	Asp	Gly	Gly	Thr 795	Trp	Glu	Gln	Gly	Gly 800
30	٧a٦	Glu	Gly	Lys	меt 805	Arg	Gly	Ser	Thr	Gln 810	Thr	val	Gly	ıle	Ala 815	Ala
	Lys	Thr	Gly	Glu 820	Asn	Thr	Thr	Ala	Ala 825	Ala	Thr	Leu	Gly	Met 830	Gly	Arg
35	Ser	Thr	Trp 835	Ser	Glu	Asn	Ser	Ala 840	Asn	Ala	Lys	Thr	Asp 845	Ser	Ile	Ser
	Leu	Phe 850	Ala	Gly	Ile	Arg	ніs 855	Asp	Ala	Gly	Asp	11e 860	Glу	Tyr	Leu	Lys
40	Gly 865	Leu	Phe	Ser	Tyr	Gly 870	Arg	Tyr	Lys	Asn	Ser 875	Ile	Ser	Arg	Ser	Thr 880
	Gly	Ala	Asp	Glu	ніs 885	Ala	Glu	Gly	Ser	Val 890	Asn	Gly	Thr	Leu	Met 895	Gln
45	Leu	Gly	Ala	Leu 900	Gly	Gly	val	Asn	va1 905	Pro	Phe	Ala	Ala	Thr 910	Gly	Asp
43	Leu	Thr	Val 915	Glu	Gly	Gly	Leu	Arg 920	Tyr	Asp	Leu	Leu	Lys 925	Gln	Asp	Ala
	Phe	Ala 930	Glu	Lys	Gly	Ser	Ala 935	Leu	Gly	Тгр	Ser	Gly 940	Asn	Ser	Leu	Thr
50	Glu 945	Gly	Thr	Leu	val	G]y 950	Leu	Ala	Gly	Leu	Lys 955	Leu	Ser	Gln	Pro	Leu 960
	Ser	Asp	Lys	Ala	Va1 965	Leu	Phe	ΑΊа	Thr	Ala 970	Gly	val	Glu	Arg	Asp 975	Leu
55	Asn	Gly	Arg	Asp 980	Tyr	Thr	٧a٦	Thr	G]y 985	Gly	Phe	Thr	Gly	Ala 990	Thr	Ala

	Ala	Thr	G]y 995	Lys	Thr	Gly	Ala	Arg 1000		Met	Pro	His	Thr 100		Leu	٧a٦
5	Ala	Gly 1010		Gly	Ala	Asp	Val 1015		Phe	Glу	Asn	Gly 1020		Asn	Gly	Leu
	Ala 1025	Arg	Tyr	Ser	Tyr	Ala 1030		Ser	Lys	Gln	Tyr 1035		Asn	His	Ser	Gly 1040
10	Arg	۷al	Gly	٧a٦	Gly 1045		Arg	Phe	Leu	Glu 1050		Gly	Gly	Gly	Thr 1055	
	Ser	Ala	Thr	Asn 1060		Asp	Asp	٧a٦	Lys 1065		Аlа	Ala	Thr	Val 1070		Ile
15	Ala	Ala	Ala 1075		Asn	Asn	Glу	G]n 1080		Ile	Asn	Glу	Phe 1085		Аlа	Gly
	Glu	Thr 1090		Tyr	Asp	Ile	Asp 1095		Asp	Gly	Thr	I]e 1100		Lys	Lys	Asp
20	Ala 1105	Thr	Ala	Ala	Asp	val 1110		Ala	Asp	Asp	Phe 1115		Gly	Leu	Gly	Leu 1120
	Lys	Lys	٧a٦	٧a٦	Thr 1125		Leu	Thr	Lys	Thr 1130		Asn	Glu	Asn	Lys 1135	
25	Asn	٧a٦	Asp	Ala 1140		٧a٦	Lys	Ala	Ala 1145		Ser	Glu	Ile	Glu 1150		Leu
	Thr	Thr	Lys 1155		Ala	Asp	Thr	Asp 1160		ΑΊа	Leu	Ala	Asp 1165		Asp	Ala
30	Ala	Leu 117(		Ala	Thr	Thr	Asn 1175		Leu	Asn	Lys	Leu 1180		Glu	Asn	Ile
	Thr 1189	Thr	Phe	Ala	Glu	Glu 1190		Lys	Thr	Asn	Ile 1199		Lys	Ile	Asp	Glu 1200
35	Lys	Leu	Glu	Ala	val 1205		Asp	Thr	val	Asp 1210		ніѕ	Аlа	Glu	Ala 1215	
	Asn	Asp	Ile	Ala 1220		Ser	Leu	Asp	Glu 1225		Asn	Thr	Lys	Ala 1230		Glu
40	Ala	۷a٦	Lys 1235		Ala	Asn	Glu	Ala 1240		Gln	Thr	Ala	Glu 1245		Thr	Lys
40	Gln	Asn 1250	val )	Asp	Ala	Lys	val 1255		Ala	ΑΊа	Glu	Thr 1260		Ala	Gly	Lys
	Ala 1265	Glu 5	Ala	Ala	Ala	Gly 1270		Аlа	Asn	Thr	Ala 1275		Asp	Lys	Ala	Glu 1280
45	Ala	۷al	Ala	Ala	Lys 1285		Thr	Asp	Ile	Lys 1290		Asp	Ile	Ala	Thr 1295	
	Lys	Asp	Asn	11e 1300		Lys	Lys	ΑΊа	Asn 1305		Ala	Asp	val	Tyr 1310		Arg
50	Glu	Glu	Ser 1315		Ser	Lys	Phe	val 1320		Ile	Asp	Gly	Leu 1325		ΑΊα	Thr
	Thr	Glu 1330		Leu	Asp	Thr	Arg 1335		Ala	Ser	Ala	G]u 1340		Ser	Ile	Ala
55	Asp 1345	His	Asp	Thr	Arg	Leu 1350		Gly	Leu	Asp	Lys 1355		٧al	Ser	Asp	Leu 1360

```
Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly
1365 1370 1375
                           Leu Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val
1380 1385 1390
                           Gly Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg
1395 1400 1405
                           Phe Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser
1410 1415 1420
10
                           Ser Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp Leu
1425 1430 1435 1446
                            Glu His His His His His
15
                                                                            1445
                            <210>
                                                   106
                                                    4179
                            <211>
                            <212>
                                                   DNA
                                                   Artificial Sequence
                            <213>
20
                            <220>
                                                   deltaG983-961c
                            <223>
                            <400>
                          atgacttctg cgcccgactt caatgcaggc ggtaccggta tcggcagcaa cagcagagca acaacagcga aatcagcagc agtatcttac gccggtatca agaacgaaat gtgcaaaagac agaagcatgc tctgtgccgg tcgggatgac gttgcggtta cagacaggga tgccaaaatc aatgccccc ccccgaatct gcataccgga gactttccaa acccaaatga cgcatacaag aatttgatca acctcaaacc tgcaattgaa gcaggctata caggacgcgg ggtagaggta ggtatcgtcg acacaggcga atccgtcggc agcatatcct ttcccgaact gtatggcaga aaagaacacg gctataacga aaattacaaa aactatacgg cgtatatgcg gaaggaagcg cctgaaggaag gaggcggtaa agcaattgaa gctctttcg acgacacat cgattagag
                                                                                                                                                                                                                                           60
                                                                                                                                                                                                                                           120
25
                                                                                                                                                                                                                                           180
                                                                                                                                                                                                                                           240
                                                                                                                                                                                                                                           300
                                                                                                                                                                                                                                           360
                                                                                                                                                                                                                                           420
                                                                                                                                                                                                                                           480
                           actgaagcaa agccgacgga tatccgccac gtaaaagaaa tcggacacat cgatttggtc
tcccatatta ttggcgggcg ttccgtggac ggcagacctg caggcggtat tgcgcccgat
gcgacgctac acataatgaa tacgaatgat gaaaccaaga acgaaatgat ggttgcagcc
30
                                                                                                                                                                                                                                           540
                                                                                                                                                                                                                                           600
                                                                                                                                                                                                                                           660
                           gegatgetat acataltyaa tacgaatgat gaaaccaaga acgaaatgat ggttgcaget atccgcaatg catgggtcaa gctggggcgaa cgtggcgtgc gcatcgtcaa taaccagttt ggaaccaacat cgagggcagg cactgccgac cttttccaaa tagccaattc ggaggagcag taccgccaag cgttgctcga ctattccggc ggtgataaaa cagacgaggg tatccgcctg atgcaacaag gcgattacgg caacctgtcc taccacatcc gtaataaaaa catgcttttc atcttttcga caggcaatga cgcacaagct cagcccaaca catatgccct attgccattt tatgaaaaaaa acgctcaaaa aggcattat acaagtcgcag gcgtagaccg cagtggagaa
                                                                                                                                                                                                                                           720
                                                                                                                                                                                                                                           780
                                                                                                                                                                                                                                           840
                                                                                                                                                                                                                                           900
35
                                                                                                                                                                                                                                           960
                                                                                                                                                                                                                                           1020
                          aagttcaaac gggaaatgta tggagaaccg ggtacagaac cgcttgagta tggctccaac cattgcggaa ttactgccat gtggtgcctg tcggcaccct atgaagcaag cgtccgtttc acccgtacaa acccgattca aattgccgga acatcctttt ccgcacccat cgtaaccgg acggcggctc tgctgctaca gaaatacccg tggatgagca acgacaacct gcgtaccacg ttgctgacga cggctcagga catcggtgca gtcggcgtgg acagcaagtt cggctgggga ctgctggatg cgggtaaggc catgaacgga cccgcgtcct ttccgttcgg cgacttacc gccgataccga aaggtacatc cgatattgcc tactccttcc gtaacgacat ttcaggcacg ggcggccga tcaaaaaagg cggcagcaa ctgcaactag acggcaacaa cacctatacg
                                                                                                                                                                                                                                           1080
                                                                                                                                                                                                                                           1140
                                                                                                                                                                                                                                           1200
                                                                                                                                                                                                                                           1260
40
                                                                                                                                                                                                                                           1320
                                                                                                                                                                                                                                           1380
                                                                                                                                                                                                                                           1440
                                                                                                                                                                                                                                           1500
                           ggcaaaacca ttatcgaagg cggttcgctg gtgttgtacg gcaacaaca atcggatatg ggcaaaacca ttatcgaagg cggttcgctg gtgttgtacg gcaacaaca atcggatatg ggcgtcgaaa ccaaaggtg gctgatttat aacggggcgg catccggcgg cagcctgaac agcgacggca ttgtctatct ggcagatacc gaccaatccg gcgcaaacga aaccgtacac atcaaaggca gtctgcagct ggacggcaaa ggtacgctgt acacacgttt gggcaaactg ctgaaagtgg acggtacggc gattatcggc ggcaagctgt acatgtcggc acgcggcaag ggggcaggct atctcaacag taccggacga cgtgttccct tcctgagtgc cgccaaaaatc
                                                                                                                                                                                                                                           1560
                                                                                                                                                                                                                                           1620
45
                                                                                                                                                                                                                                           1680
                                                                                                                                                                                                                                           1740
                                                                                                                                                                                                                                           1800
                                                                                                                                                                                                                                           1860
                           gggcaggatt attettett cacaaacate gaaaccgacg gcggcctgct ggcttcctc gacagcgtcg aaaaaacagc gggcagtgaa ggcgacacgc tgtcctatta tgtccgtcgc ggcaatgcgg cacggactgc ttcggcagcg gcacattccg cgcccgccgg tctgaaacac
                                                                                                                                                                                                                                           1920
                                                                                                                                                                                                                                           1980
                                                                                                                                                                                                                                           2040
50
                                                                                                                                                                                                                                           2100
                            gccgtagaac agggcggcag caatctggaa aacctgatgg tcgaactgga tgcctccgaa
                           tcatccgcaa cacccgagac ggttgaaact gcggcagccg accgcacaga tatgccgggc atccgccct acggcgcaac tttccgcca gcggcagccg tacagcatgc gaatgccgcc gacggtgtac gcatcttcaa cagtctcgcc gctaccgtc tatgccgacag taccgccccatgccgata tgcagggacg ccgcctgaaa gccgtatcgg acgggttgga ccacaacggcacgggtctgc gcgtcatcgc gcaaacccaa caggacggtg gaacgtggga acagggcggtgtgaaggca aaatgcgcgg cagtacccaa accgtcggca ttgccgcgaa aaccggcgaa
                                                                                                                                                                                                                                           2160
                                                                                                                                                                                                                                           2220
2280
                                                                                                                                                                                                                                           2340
                                                                                                                                                                                                                                           2400
55
                                                                                                                                                                                                                                           2460
                            aatacgacag cagccgccac actgggcatg ggacgcagca catggagcga aaacagtgca
                                                                                                                                                                                                                                           2520
```

2580

2640 2700

2760

3180 3240 3300

3360

3420 3480 3540

3600 3660 3720

3780 3840 3900

3960 4020

4080

4140 4179

```
aatgcaaaaa ccgacagcat tagtctgttt gcaggcatac ggcacgatgc gggcgatatc
             ggctatctca aaggcctgtt ctcctacgga cgctacaaaa acagcatcag ccgcagcacc
ggtgcggacg aacatgcgga aggcagcgtc aacggcacgc tgatgcagct gggcgcactg
             ggcggtgtca acgttccgtt tgccgcaacg ggagatttga cggtcgaagg cggtctgcgc
             aggregated acgreecest typegradery graduating egyptergady egyptergege tacgacetre teaaacagga tycattegee gaaaaaggca gtgetttygg etyggagegge aacageetra etygaaggca getggtegga etegggget tyaageetyte geaaceetry agegataaag eegteetyt tyeaacggeg gyegtggaae gegaeetyaa egygaeggae tacaeggtaa egygegget taceggegg aetygeagea eeggeaagae gygggeaege aatatgeege acaeeegget gygtaeege etygaaegget tygeaegtta eagetaegee gytteeaaaa agtaeeggeaa eeaaeggga
             cgagtcggcg taggctaccg gttcctcgag ggtggcggag gcactggatc cgccacaaac gacgacgatg ttaaaaaagc tgccactgtg gccattgctg ctgcctacaa caatggccaa gaaatcaacg gtttcaaagc tggagagacc atctacgaca ttgatgaaga cggcacaatt
10
             accaaaaaag acgcaactgc agccgatgtt gaagccgacg actttaaagg tctgggtctg
             aaaaaagtcg tgactaacct gaccaaaacc gtcaatgaaa acaaacaaaa cgtcgatgcc
aaagtaaaag ctgcagaatc tgaaatagaa aagttaacaa ccaagttagc agacactgat
             gccgctttag cagatactga tgccgctctg gatgcaacca ccaacgcctt gaataaattg
             ggagaaaata taacgacatt tgctgaagag actaagacaa atatcgtaaa aattgatgaa aaattagaag ccgtggctga taccgtcgac aagcatgccg aagcattcaa cgatatcgcc
             gattcattgg atgaaaccaa cactaaggca gacgaagccg tcaaaaccgc caatgaagcc
             aaacagacgg ccgaagaaac caaacaaaac gtcgatgcca aagtaaaagc tgcagaaact gcagcaggca aagccgaagc tgccgctggc acagctaata ctgcagccga caagccgaa
             gctgtcgctg caaaagttac cgacatcaaa gctgatatcg ctacgaacaa agataatatt
20
             gctaaaaaag caaacagtgc cgacgtgtac accagagaag agtctgacag caaatttgtc
agaattgatg gtctgaacgc tactaccgaa aaattggaca cacgcttggc ttctgctgaa
             aāatccāttō ccgatcacōa tactcgcctg aacggtttgg ataāaacāgt gtcaōacctg
             cgcaaagaaa cccgccaagg ccttgcagaa caagccgcgc tctccggtct gttccaacct
tacaacgtgg gtctcgagca ccaccaccac caccactga
25
             <210>
                         107
                         1392
             <211>
             <212>
                         PRT
                         Artificial Sequence
             <213>
             <220>
             <223>
                         deltaG983-961c
30
             Met Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser 1 \hspace{1cm} 5 \hspace{1cm} 15
             Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly
20 25 30
35
             Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro 50 60
40
             Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys 70 75 80
             Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg
85 90 95
45
             Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile 100 \hspace{1cm} 105
             Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn
115 120
50
             Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly 130 140
             Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu
145 150 155
55
             Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His
```

					TP2					1/0					1/5	
5	Ile	Asp	Leu	Val 180	Ser	His	Ile	Ile	Gly 185	Gly	Arg	Ser	۷al	Asp 190	Gly	Arg
	Pro	Ala	Gly 195	Gly	Ile	Ala	Pro	Asp 200	Ala	Thr	Leu	His	11e 205	Met	Asn	Thr
10	Asn	Asp 210	Glu	Thr	Lys	Asn	Glu 215	Met	Met	val	Аlа	Ala 220	Ile	Arg	Asn	Ala
	Trp 225	Val	Lys	Leu	Gly	G1u 230	Arg	Gly	۷al	Arg	11e 235	Val	Asn	Asn	Ser	Phe 240
15	Gly	Thr	Thr	Ser	Arg 245	Ala	Gly	Thr	Ala	Asp 250	Leu	Phe	Gln	Ile	Ala 255	Asn
	Ser	Glu	Glu	G]n 260	Tyr	Arg	Gln	Аlа	Leu 265	Leu	Asp	Tyr	ser	Gly 270	Gly	Asp
00	Lys	Thr	Asp 275	Glu	Gly	Ile	Arg	Leu 280	Met	Gln	Gln	Ser	Asp 285	Tyr	Gly	Asn
20	Leu	ser 290	Tyr	His	Ile	Arg	Asn 295	Lys	Asn	Met	Leu	Phe 300	Ile	Phe	Ser	Thr
	G]y 305	Asn	Asp	Ala	Gln	Ala 310	Gln	Pro	Asn	Thr	Tyr 315	Аlа	Leu	Leu	Pro	Phe 320
25	Tyr	Glu	Lys	Asp	Ala 325	Gln	Lys	Gly	Ile	11e 330	Thr	val	Ala	Gly	va1 335	Asp
	Arg	Ser	Gly	Glu 340	Lys	Phe	Lys	Arg	Glu 345	Met	Tyr	Glу	Glu	Pro 350	Gly	Thr
30	Glu	Pro	Leu 355	Glu	Tyr	Gly	Ser	Asn 360	His	Cys	Gly	Ile	Thr 365	Ala	Met	Тгр
	Cys	Leu 370	Ser	Ala	Pro	Tyr	G1u 375	Ala	Ser	val	Arg	Phe 380	Thr	Arg	Thr	Asn
35	Pro 385	ıle	Gln	Ile	Ala	G]y 390	Thr	Ser	Phe	ser	Ala 395	Pro	ıJe	val	Thr	Gly 400
	Thr	Аla	Ala	Leu	Le <b>u</b> 405	Leu	Gln	Lys	Tyr	Pro 410	Trp	Met	Ser	Asn	Asp 415	Asn
40	Leu	Arg	Thr	Thr 420	Leu	Leu	Thr	Thr	Ala 425	Gln	Asp	Ile	Gly	Ala 430	٧a٦	Gly
	Val	Asp	Ser 435		Phe	Gly		Gly 440		Leu		Ala		Lys	Ala	Met
45	Asn	Gly 450	Pro	Ala	Ser	Phe	Pro 455	Phe	Gly	Asp	Phe	Thr 460	Ala	Asp	Thr	Lys
	Gly 465	Thr	Ser	Asp	Ile	Ala 470	Tyr	ser	Phe	Arg	Asn 475	Asp	Ile	Ser	Gly	Thr 480
50	G∃y	Gly	Leu	Ile	Lys 485	Lys	Gly	Gly	Ser	Gln 490	Leu	Gln	Leu	His	Gly 495	Asn
	Asn	Thr	Tyr	Thr 500	Gly	Lys	Thr	Ile	11e 505	Glu	Gly	Gly	Ser	Leu 510	٧a٦	Leu
55	Tyr	Gly	Asn 515	Asn	Lys	ser	Asp	Met 520	Arg	۷al	Glu	Thr	Lys 525	Glу	ΑΊа	Leu
	Ile	Tyr	Asn	Gly	Ala	Ala	Ser	GТу	Gly	Ser	Leu	Asn	ser	Asp	Gly	ıle

530 535 Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His 545 550 555 560 Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg 565 570 575 Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys 580 585 590 10 Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr 595 600 605 Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr 610 615 620 15 Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala Ser Leu 625 630 635 Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu Ser Tyr 645 650 655 20 Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala Ala His 660 665 670 Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln Gly Gly Ser Asn 675 680 685 25 Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu Ser Ser Ala Thr 690 695 700 Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr Asp Met Pro Gly 705 710 715 Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Val Gln His 725 730 735 30 Ala Asn Ala Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala Ala Thr 740 745 750 Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly Arg Arg 755 760 765 35 Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly Leu Arg 770 775 780 Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln Gly Gly 785 790 795 800 40 Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala 805 810 Lys Thr Gly Glu Asn Thr Thr Ala Ala Ala Thr Leu Gly Met Gly Arg 820 825 830 45 Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser 835 840 845 Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys 850 855 50 Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr 865 870 875 880 Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln 885 890 895 55 Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp

		900	905		910
5	Leu Thr Val 915	Glu Gly Gly	Leu Arg Tyr 920	Asp Leu Leu Lys 925	Gln Asp Ala
	Phe Ala Glu 930	Lys Gly Ser	Ala Leu Gly 935	Trp Ser Gly Asn 940	Ser Leu Thr
10	Glu Gly Thr 945	Leu Val Gly 950	Leu Ala Gly	Leu Lys Leu Ser 955	Gln Pro Leu 960
	Ser Asp Lys	Ala Val Leu 965	Phe Ala Thr	Ala Gly Val Glu 970	Arg Asp Leu 975
15	Asn Gly Arg	Asp Tyr Thr 980	Val Thr Gly 985	Gly Phe Thr Gly	Ala Thr Ala 990
13	Ala Thr Gly 995	Lys Thr Gly	Ala Arg Asn 1000	Met Pro His Thr	
	Ala Gly Leu 1010	Gly Ala Asp	val Glu Phe 1015	Gly Asn Gly Trp 1020	Asn Gly Leu
20	Ala Arg Tyr 1025	Ser Tyr Ala 1030		Gln Tyr Gly Asn 1035	His Ser Gly 1040
	Arg Val Gly	Val Gly Tyr 1045	Arg Phe Leu	Glu Gly Gly Gly 1050	Gly Thr Gly 1055
25	Ser Ala Thr	Asn Asp Asp 1060	Asp Val Lys 1065	Lys Ala Ala Thr	Val Ala Ile 1070
	Ala Ala Ala 107		Gly Gln Glu 1080	Ile Asn Gly Phe 108	
30	Glu Thr Ile 1090	Tyr Asp Ile	Asp Glu Asp 1095	Gly Thr Ile Thr 1100	Lys Lys Asp
	Ala Thr Ala 1105	Ala Asp Val		Asp Phe Lys Gly 1115	Leu Gly Leu 1120
35	Lys Lys Val	Val Thr Asn 1125	Leu Thr Lys	Thr Val Asn Glu 1130	Asn Lys Gln 1135
	Asn Val Asp	Ala Lys Val 1140	Lys Ala Ala 1145	Glu Ser Glu Ile	Glu Lys Leu 1150
40	Thr Thr Lys 115		Thr Asp Ala 1160	Ala Leu Ala Asp 116	
	Ala Leu Asp 1170	Ala Thr Thr	Asn Ala Leu 1175	Asn Lys Leu Gly 1180	Glu Asn Ile
45	Thr Thr Phe 1185	Ala Glu Glu 1190		Asn Ile Val Lys 1195	Ile Asp Glu 1200
	Lys Leu Glu	Ala Val Ala 1205	Asp Thr Val	Asp Lys His Ala 1210	Glu Ala Phe 1215
50	Asn Asp Ile	Ala Asp Ser 1220	Leu Asp Glu 1225	Thr Asn Thr Lys	Ala Asp Glu 1230
	Ala Val Lys 123		Glu Ala Lys 1240	Gln Thr Ala Glu 124	
55	Gln Asn Val 1250	Asp Ala Lys	Val Lys Ala 1255	Ala Glu Thr Ala 1260	Ala Gly Lys
	Ala Glu Ala	Ala Ala Gly	Thr Ala Asn	Thr Ala Ala Asp	Lys Ala Glu

```
1270
                     1265
                                                                                                             1275
                                                                                                                                                          1280
                     Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn 1285 1290 1295
                     Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg
1300 1305 1310
                     Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr
1315 1320 1325
10
                     Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala
1330 1335 1340
                     Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu 1345 1350 1355 136
15
                     Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly
1365 1370 1375
                     Leu Phe Gln Pro Tyr Asn Val Gly Leu Glu His His His His His His 1380 1385 1390
20
                     <210>
                                      108
                     <211>
                                      1947
                      <212>
                                      DNA
                      <213>
                                      Artificial Sequence
25
                     <220>
                                      deltaG741-961
                     <223>
                     <400>
                                      108
                     atggtcgccg ccgacatcgg tgcggggctt gccgatgcac taaccgcacc gctcgaccat
                                                                                                                                                                            60
                     aaagacaaag gtttgcagtc tttgacgctg gatcagtccg tcaggaaaaa cgagaaactg aagctggcgg cacaaggtgc ggaaaaaact tatggaaacg gtgacagcct caatacgggc aaattgaaga acgacaaggt cagccgtttc gactttatcc gccaaatcga agtggacggg cagctcatta ccttggagag tggagagttc caagtataca aacaaagcca ttccgcctta
                                                                                                                                                                            120
30
                                                                                                                                                                            180
                                                                                                                                                                            240
                                                                                                                                                                            300
                     accgcctttc agaccgagca aatacaagat tcggagcatt ccgggaagat ggttgcgaaa
                                                                                                                                                                            360
                    accgcctttc agaccgagca aatacaagat tcggagcatt ccgggaagat ggttgcgaaa cgccagttca gaatcggcga catagcgggc gaacatacat cttttgacaa gcttcccgaa ggcggcaggg cgacatatcg cgggaaggcg ttcggttcag acgatgccgg cggaaaactg acctacacca tagatttcgc cgccaagcag ggaaacggca aaatcgaaca tttgaaatcg ccagaactca atgtcgacct ggccgccgc gatatcaagc cggatggaaa acgccatgcc gtcatcagcg gttccgtcct ttacaaccaa gccgagaaag gcagttactc cctcggtatc tttggcggaa aagcccagga agttgccgc agcgggaagg tgaaaaccgt aaacggcata cgccatatcg gcccttgccgc caagcaactc gagggtggcg gaggcactgg atccgccaca
                                                                                                                                                                            420
                                                                                                                                                                            480
35
                                                                                                                                                                            540
                                                                                                                                                                            600
                                                                                                                                                                            660
                                                                                                                                                                            720
                                                                                                                                                                            780
                     aacgacgacg atgttaaaaa agctgccact gtggccattg ctgctgccta caacaatggc
caagaaatca acggtttcaa agctggagag accatctacg acattgatga agacggcaca
attaccaaaa aagacgcaac tgcagccgat gttgaagccg acgactttaa aggtctgggt
                                                                                                                                                                            840
                                                                                                                                                                            900
40
                                                                                                                                                                            960
                     ctgaaaaaag tcgtgactaa cctgaccaaa accgtcaatg aaaacaaaca aaacgtcgat
                                                                                                                                                                            1020
                     gccaaagtaa aagctgcaga atctgaaata gaaaagttaa caaccaagtt agcagacact
gatgccgctt tagcagatac tgatgccgct ctggatgcaa ccaccaacgc cttgaataaa
ttgggagaaa atataacgac atttgctgaa gagactaaga caaatatcgt aaaaattgat
                                                                                                                                                                            1080
                                                                                                                                                                            1140
                                                                                                                                                                            1200
                     gaaaaattag aagccgtggc tgataccgtc gacaagcatg ccgaagcatt caacgatatc
gccgattcat tggatgaaac caacactaag gcagacgaag ccgtcaaaac cgccaatgaa
gccaaacaga cggccgaaga aaccaaacaa aacgtcgatg ccaaagtaaa agctgcagaa
                                                                                                                                                                            1260
45
                                                                                                                                                                            1320
                                                                                                                                                                            1380
                    actgcagcag gcaaagccga agctgccgct ggcacagcta atactgcagc cgacaaggcc
gaagctgtcg ctgcaaaagt taccgacatc aaagctgata tcgctacgaa caaagataat
attgctaaaa aagcaaacag tgccgacgtg tacaccagag aagagtctga cagcaaattt
gtcagaattg atggtctgaa cgctactacc gaaaaattgg acacacgctt ggcttctgct
gaaaaatcca ttgccgatca cgatactcgc ctgaacggtt tggataaaac agtgtcagac
ctgcgcaaag aaacccgcca aggccttgca gaacaagccg cgctctccgg tctgttccaa
ccttacaacg tgggtcggtt caatgtaacg gctgcagtcg gcggctacaa atccgaatcg
gcagtcgcca tcggtaccgg ttctgcaga accaacact ttgccgcaa agcaggcgtg
                                                                                                                                                                            1440
                                                                                                                                                                            1500
                                                                                                                                                                            1560
                                                                                                                                                                            1620
                                                                                                                                                                            1680
50
                                                                                                                                                                            1740
                                                                                                                                                                            1800
                                                                                                                                                                            1860
                     gcagteggea ettegteegg ttetteegea geetaecatg teggegteaa ttaegagtgg
                                                                                                                                                                            1920
                     ctcgagcacc accaccacca ccactga
                                                                                                                                                                            1947
55
                     <210>
                                      109
                     <211>
                                      648
```

	<212 <213		PRT Arti	ifici	ial s	seque	ence									
5	<220 <223		delt	aG74	11-96	51										
	<400 Met 1		109 Ala	Αla	Asp 5	Ile	Gly	Ala	Gly	Leu 10	Αla	Asp	Ala	Leu	Thr 15	Ala
10	Pro	Leu	Asp	His 20	Lys	Asp	Lys	GТу	Leu 25	Gln	Ser	Leu	Thr	Leu 30	Asp	Gln
	ser	val	Arg 35	Lys	Asn	Glu	Lys	Leu 40	Lys	Leu	Δla	Аlа	G]n 45	Gly	Ala	Glu
15	Lys	Thr 50	Tyr	Glу	Asn	Gly	Asp 55	Ser	Leu	Asn	Thr	G]y 60	Lys	Leu	Lys	Asn
	Asp 65	Lys	val	Ser	Arg	Phe 70	Asp	Phe	Ile	Arg	G]n 75	Ile	Glu	۷a٦	Asp	G]y 80
20	Gln	Leu	Ile	Thr	Leu 85	Glu	Ser	Gly	Glu	Phe 90	Gln	val	Tyr	Lys	G]n 95	Ser
	His	Ser	Ala	Leu 100	Thr	Ala	Phe	Gln	Thr 105	Glu	Gln	Ile	Gln	Asp 110	Ser	Glu
25	His	Ser	Gly 115	Lys	Met	٧a٦	Ala	Lys 120	Arg	Gln	Phe	Arg	Ile 125	Gly	Asp	IJе
	Ala	Gly 130	Glu	His	Thr	ser	Phe 135	Asp	Lys	Leu	Pro	Glu 140	Gly	Gly	Arg	Аla
30	Thr 145	Туг	Arg	Gly	Thr	Ala 150	Phe	Gly	Ser	Asp	Asp 155	Аlа	Gly	Gly	Lys	Leu 160
	Thr	Tyr	Thr	Ile	Asp 165	Phe	Ala	Ala	Lys	Gln 170	Gly	Asn	Gly	Lys	11e 175	Glu
35	His	Leu	Lys	ser 180	Pro	Glu	Leu	Asn	∨a1 185	Asp	Leu	Ala	Ala	аlа 190	Asp	Ile
	Lys	Pro	Asp 195	Gly	Lys	Arg	His	Ala 200	∨al	Ile	Ser	Gly	Ser 205	∨al	Leu	Tyr
40	Asn	G]n 210	Ala	Glu	Lys	Gly	Ser 215	Tyr	Ser	Leu	Gly	11e 220	Phe	Gly	Gly	Lys
	Ala 225	Gln	Glu	val	Ala	Gly 230	Ser	Ala	Glu	Val	Lys 235	Thr	٧a٦	Asn	Gly	11e 240
45	Arg	His	Ile	Gly	Leu 245	Ala	Ala	Lys	Gln	Leu 250	Glu	Gly	Gly	Gly	Gly 255	Thr
	Gly	ser	Ala	Thr 260	Asn	Asp	Asp	Asp	va1 265	Lys	Lys	Ala	Ala	Thr 270	val	ΑΊа
50	Ile	Ala	Ala 275	Ala	Tyr	Asn	Asn	Gly 280	Gln	Glu	Ile	Asn	Gly 285	Phe	Lys	Ala
30	Gly	Glu 290	Thr	Ile	Tyr	Asp	11e 295	Asp	Glu	Asp	Gly	Thr 300	Ile	Thr	Lys	Lys
	Asp 305	Ala	Thr	Ala	Ala	Asp 310	val	Glu	Ala	Asp	Asp 315	Phe	Lys	Gly	Leu	G1y 320
55	Leu	Lys	Lys	val	Va1 325	Thr	Asn	Leu	Thr	Lys 330	Thr	val	Asn	Glu	Asn 335	Lys

```
Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys
340 345 350
             Leu Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp 355 360 365
            Ala Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn 370 380
             Ile Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp
385 390 395 400
10
             Glu Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala
405 410 415
            Phe Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp 420 425 430
15
             Glu Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr
435 440 445
             Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly
450 455 460
20
             Lys Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala
465 470 475 480
             Glu Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr
485 490 495
25
             Asn Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr 500 505 510
             Arg Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala
515 520 525
30
             Thr Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile
530 535 540
             Ala Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp 545 550 555
35
             Leu Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser
565 570 575
             Gly Leu Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala
580 585 590
40
             Val Gly Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe
595 600 605
             Arg Phe Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr 610 620
45
             Ser Ser Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp 625 630 635
             Leu Glu His His His His His His 645
50
                       110
             <210>
             <211>
                       1782
             <212>
                       DNA
                       Artificial Sequence
             <213>
             <220>
55
             <223>
                       deltaG741-961c
```

```
<400>
                        110
            atggtcgccg ccgacatcgg tgcggggctt gccgatgcac taaccgcacc gctcgaccat
                                                                                                             60
                                                                                                             120
                                                                                                             180
                                                                                                              240
                                                                                                              300
                                                                                                             360
420
                                                                                                              480
                                                                                                             540
10
                                                                                                             600
                                                                                                              660
                                                                                                             720
780
                                                                                                             840
                                                                                                             900
15
                                                                                                              960
                                                                                                              1020
             gccaaagtaa aagctgcaga atctgaaata gaaaagttaa caaccaagtt agcagacact
gatgccgctt tagcagatac tgatgccgct ctggatgcaa ccaccaacgc cttgaataaa
                                                                                                             1080
                                                                                                             1140
            ttgggagaaa atataacgac atttgctgaa gagactaaga caaatatcgt aaaaattgat gaaaaattag aagccgtggc tgataccgtc gacaagcatg ccgaagcatt caacgatatc gccgattcat tgggtgaaac caacactaag gcagacgaag ccgtcaaaac cgccaatgaa gccaaacaa aacgtcgatg ccaaagtaaa agctgcagaag
                                                                                                              1200
                                                                                                             1260
                                                                                                             1320
20
                                                                                                              1380
             actgcagcag gcaaagccga agctgccgct ggcacagcta atactgcagc cgacaaggcc
gaagctgtcg ctgcaaaagt taccgacatc aaagctgata tcgctacgaa caaagataat
                                                                                                             1440
                                                                                                             1500
            attgctaaaa aagcaaacag tgccgacgtg tacaccagag aagagtctga cagcaaattt gtcagaattg atggtctgaa cgctactacc gaaaaattgg acacacgctt ggcttctgct gaaaaatcca ttgccgatca cgatactcgc ctgaacggtt tggataaaac agtgtcagac
                                                                                                             1560
                                                                                                             1620
1680
25
             ctgcgcaaag aaacccgcca aggccttgca gaacaagccg cgctctccgg tctgttccaa
                                                                                                              1740
             ccttacaacg tgggtctcga gcaccaccac caccaccact ga
                                                                                                              1782
             <210>
             <211>
                        593
             <212>
                        PRT
30
                        Artificial Sequence
             <213>
             <220>
                        deltaG741-961c
             <223>
             Met Val Ala Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala
1 5 10 15
35
             Pro Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln 20 25 30
             Ser Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln Gly Ala Glu 35 40 45
40
             Lys Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys Leu Lys Asn 50 55 60
             Asp Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu Val Asp Gly 65 70 75 80
45
             Gln Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr Lys Gln Ser
85 90 95
            His Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln Asp Ser Glu 100 \hspace{1cm} 105 \hspace{1cm} 110
50
            His Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile Gly Asp Ile
115 120 125
            Ala Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala
130 135 140
55
             Thr Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu
```

	145				150					155					160
5	Thr Ty	r Thr	Ile	Asp 165	Phe	Ala	Ala	Lys	G]n 170	Gly	Asn	Gly	Lys	17e 175	Glu
	His Le	u Lys	Ser 180	Pro	Glu	Leu	Asn	Val 185	Asp	Leu	Ala	Ala	Ala 190	Asp	Ile
10	Lys Pr	o Asp 195	Gly	Lys	Arg	His	Ala 200	٧al	Ile	Ser	Gly	Ser 205	۷al	Leu	Tyr
	Asn Gl 21		Glu	Lys	Gly	Ser 215	Tyr	Ser	Leu	Gly	11e 220	Phe	Gly	Gly	Lys
15	Ala Gl 225	n Glu	val	Ala	G]y 230	ser	Ala	Glu	val	Lys 235	Thr	val	Asn	Gly	11e 240
	Arg Hi	s Ile	Gly	Leu 245	Ala	Ala	Lys	Gln	Leu 250	Glu	Gly	Gly	Gly	G]y 255	Thr
20	Gly Se	r Ala	Thr 260	Asn	Asp	Asp	Asp	va1 265	Lys	Lys	Ala	Ala	Thr 270	val	Ala
	Ile Al	a Ala 275	Ala	Tyr	Asn	Asn	Gly 280	Gln	Glu	Ile	Asn	Gly 285	Phe	Lys	Ala
25	Gly Gl 29		Ile	Tyr	Asp	11e 295	Asp	Glu	Asp	Gly	Thr 300	Ile	Thr	Lys	Lys
	Asp Al 305	a Thr	Ala	Ala	Asp 310	val	Glu	Ala	Asp	Asp 315	Phe	Lys	Gly	Leu	G]y 320
	Leu Ly	s Lys	Val	Va1 325	Thr	Asn	Leu	Thr	Lys 330	Thr	val	Asn	Glu	Asn 335	Lys
30	Gln As	n Val	Asp 340	Ala	Lys	Val	Lys	Ala 345	Ala	Glu	Ser	Glu	11e 350	Glu	Lys
	Leu Th	r Thr 355	Lys	Leu	Ala	Asp	Thr 360	Asp	Ala	Ala	Leu	Ala 365	Asp	Thr	Asp
35	Ala Al 37	_	Asp	Ala	Thr	Thr 375	Asn	Ala	Leu	Asn	Lys 380	Leu	Gly	Glu	Asn
	Ile Th 385	r Thr	Phe	Ala	G]u 390	Glu	Thr	Lys	Thr	Asn 395	Ile	Val	Lys	Ile	400
40	Glu Ly	s Leu	Glu	Ala 405	val	Ala	Asp	Thr	val 410	Asp	Lys	His	Ala	Glu 415	Ala
	Phe As		420					425					430		
45	Glu Al	435					440					445			
	Lys Gl 45	0				455					460				
50	Lys Al 465	a Glu	Ala	Ala	470	Gly	Thr	Ala	Asn	Thr 475	Ala	Ala	Asp	Lys	480
	Glu Al			485					490					495	
55	Asn Ly	•	500					505					510		
	Arg Gl	u Glu	Ser	Asp	Ser	Lys	Phe	val	Arg	Ile	Asp	Gly	Leu	Asn	Ala

		515		5	20			525		
	Thr Thr 530	Glu Lys	Leu As	p Thr A 535	rg Lei	u Ala S	er Ala 540	Glu Lys	s Ser Ile	
5	Ala Asp 545	ніs Asp	Thr Ar		sn Gly		sp Lys	Thr Va	Ser Asp 560	
	Leu Arg	Lys Glu	Thr Ar 565	g Gln G	aly Lei	u Ala G 570	lu Gln	Ala Ala	Leu Ser 575	
10	Gly Leu	Phe Gln 580	Pro Ty	r Asn V	/al Gly 58		ilu His	His His	s His His )	
	ніѕ									
15	<210> <211> <212> <213>	112 3939 DNA Artific	ial Seq	uence						
20	<220> <223>	deltaG74	41-983							
25	aaagacaa aagctgge aaattgaa cagctca	aag gtttg cgg cacaa aga acgad tta ccttg	gcagtc aggtgc caaggt ggagag	tttgācg ggaaaaa cagccgt tggagag	ictg ga lact ta lttc ga lttc ca	atcagtc atggaaa actttat aagtata	cg tca cg gtg cc gcc ca aac	ggaaaaa acagcct aaatcga aaagcca	gctcgaccat cgagaaactg caatacgggc agtggacggg ttccgcctta ggttgcgaaa	60 120 180 240 300 360
30	cgccagte ggcggcag acctacae ccagaact gtcatcag tttggcgg	tca gaato ggg cgaca cca tagat tca atgto gcg gttco gaa aagco	cggcga atatcg tttcgc cgacct cgtcct ccagga	catagcg cgggacg cgccaag ggccgcc ttacaac agttgcc	gggc ga ggcg ti gcag ga gcc ga ccaa ga ggc aa	aacatac tcggttc gaaacgg atatcaa ccgagaa gcgcgga	at ctt ag acga ca aaa gc cgga ag gca ag tga	ttgacaa atgccgg tcgaaca atggaaa gttactc aaaccgt	gcttcccgaa cggaaaactg tttgaaatcg acgccatgcc cctcggtatc aaacggcata cacttctgcg	420 480 540 600 660 720 780
35	tcagcage tgtgccge ccgaatc ctcaaace acaggcga	tca atgca cag tatch gtc gggan tgc ataco ctg caath aat ccgto	aggcgg ttacgc tgacgt cggaga tgaagc cggcag	taccggt cggtatc tgcggtt ctttcca aggctat catatcc	tato go	gcagcaa acgaaat acaggga caaatga gacgcgg ccgaact	ca gca gt gca tg cca cg cat gg tag gt atg	gagcaac aagacag aaatcaa acaagaa aggtagg gcagaaa	aacagcgaaa aagcatgctc tgccccccc tttgatcaac tatcgtcgac agaacacggc	840 900 960 1020 1080 1140
40	ggcggtaa ccgacgga ggcgggcg ataatgaa tgggtcaa agggcagg	aag acati ata teego gtt eegto ata egaai age tgggo gea etgeo	tgaagc ccacgt ggacgg tgatga cgaacg cgacct	ttctttc aaaagaa cagacct aaccaag tggcgtg tttccaa	igac ga latc ga gca ga gaac ga gcgc a lata ga	atgaggo gacacat gcggtat aaatgat tcgtcaa ccaattc	cg tta cg att tg cgc gg ttg ta aca gg agg	tagagac tggtctc ccgatgc cagccat gttttgg agcagta	tgaagacgga tgaagcaaag ccatattatt gacgctacac ccgcaatgca aacaacatcg ccgccaagcg	1200 1260 1320 1380 1440 1500 1560
45	gattacgo ggcaatga gctcaaaa	gca accto acg cacaa aag gcati	gtčcta agctca tatcac	ccacato gcccaac agtcgca	cgt a aca ta aggc g	ataaaaa atgccct tagaccg	ca tgc at tgc ca gtg	ttttčat catttta gagaaaa	gcaacagagc cttttcgaca tgaaaaagac gttcaaacgg ttgcggaatt	1620 1680 1740 1800 1860
50	actgccat ccgattca ctgctgca gctcagga ggtaagga ggtacata aaaaaaga	tgť ggťgo aaa ttgco aga aatao aca tcggo cca tgaao ccg atato gcg gcago	cctgtc cggaac cccgtg tgcagt cggacc tgccta ccaact	ggcacco atccttt gatgago cggcgtg cgcgtco ctcctto gcaactg	tat gatter gatter gate gate gate accept accept accept gate gate gate gate gate gate gate gat	aagcaag cacccat acaacct gcaagtt cgttcgg acgacat gcaacaa	cg tcc cg taa gc gta cg gct cg act tt cag ca cct	gtttcac ccggcac ccacgtt ggggact ttaccgc gcacggg atacggg	ccgtacaaac ggcggctctg gctgacgacg gctggatgcg cgatacgaaa cggcctgatc caaaaccatt	1920 1980 2040 2100 2160 2220 2280
55	aaaggtgd gtctatcd ctgcagcd	cgc tgati tgg cagai tgg acgg	ttatāa taccga caaagg	cggggcg ccaatcc tacgctg	ggca to ggc go gtac a	ccggcgg caaacga cacgttt	ca gcc aa ccg gg gca	tgaacag tacacat aactgct	cgtcgaaacc cgacggcatt caaaggcagt gaaagtggac ggcaggctat	2340 2400 2460 2520 2580

2640

```
ctcaacagta ccggacgacg tgttcccttc ctgagtgccg ccaaaatcgg gcaggattat
tctttcttca caaacatcga aaccgacggc ggcctgctgg cttccctcga cagcgtcgaa
aaaacagcgg gcagtgaagg cgacacgctg tcctattatg tccgtcgcgg caatgcggca
                                                                                                                              2700
                                                                                                                              2760
             cggactgctt cggcagcggc acattccgcg cccgccggtc tgaaacacgc cgtagaacag ggcggcagca atctggaaaa cctgatggtc gaactggatg ctccgaatc atccgcaacac cccgagacgg ttgaaactgc ggcagccgac cgcagata tgccgggcat ccgcccctac
                                                                                                                              2820
                                                                                                                              2880
                                                                                                                              2940
             ggcgcaactt tccgcgcagc ggcagccgta cagcatgcga atgccgcga cggtgtacgc atcttcaaca gtctcgccgc taccgtctat gccgacagta ccgccgccca tgccgatatgcaggacgcc gcctgaaagc cgtatcggac gggttggacc acaacggcac gggtctgcgc
                                                                                                                              3000
                                                                                                                              3060
                                                                                                                              3120
             gtčátcgčgc áaacčcaača ggacggťgga áčgtgýgaac agggcggtgt ťgaaggcaaa
                                                                                                                              3180
             atgcgcggca gtacccaaac cgtcggcatt gccgcgaaaa ccggcgaaaa tacgacagca gccgccacac tgggcatggg acgcagcaca tggagcgaaa acagtgcaaa tgcaaaaacc gacagcatta gtctgtttgc aggcatacgg cacgatgcgg gcgatatcgg ctatctcaaa
10
                                                                                                                              3240
                                                                                                                              3300
                                                                                                                              3360
             ggccīgttct cctacggacg cīācaaaaāc agcātcāgcc ģcāgcaccģģ tgcggacgaa
                                                                                                                              3420
             catgcggaag gcagcgtcaa cggcacgctg atgcagctgg gcgcactggg cggtgtcaac gttccgtttg ccgcaacggg agatttgacg gtcgaaggcg gtctgcgcta cgacctgctc aaacaggatg cattcgccga aaaaggcagt gctttgggct ggagcggcaa cagcctcact
                                                                                                                              3480
                                                                                                                              3540
                                                                                                                              3600
15
             gaaggcacgc tggtcggact cgcgggtctg aagctgtcgc aacccttgag cgataaagcc
gtcctgtttg caacggcggg cgtggaacgc gacctgaacg gacgcgacta cacggtaacg
                                                                                                                              3660
                                                                                                                              3720
             ggcggcttta ccggcgcgac tgcagcaacc ggcaagacgg gggcacgcaa tatgccgcac acccgtctgg ttgccggct gggcgcggat gtcgaattcg gcaacggctg gaacggcttg gcacgttaca gctacgccgg ttccaaacag tacggcaacc acagcggacg agtcggcgta
                                                                                                                              3780
                                                                                                                              3840
                                                                                                                              3900
             ggctaccggt tcctcgagca ccaccaccac caccactga
                                                                                                                              3939
20
             <210>
                           1312
             <211>
             <212>
                          PRT
              <213>
                          Artificial Sequence
25
              <220>
                           deltaG741-983
             <223>
             <400>
             Met Val Ala Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala
1 5 10 15
30
             Pro Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln 20 25 30
             Ser Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln Gly Ala Glu
35 40 45
             Lys Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys Leu Lys Asn 50 55
             Asp Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu Val Asp Gly 65 75 80
40
             Gln Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr Lys Gln Ser
85 90 95
             His Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln Asp Ser Glu
100 105 110
             His Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile Gly Asp Ile
115 120 125
             Ala Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala
130 140
             Thr Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu
145 150 155 160
50
             Thr Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Ile Glu
165 170 175
             His Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala Asp Ile
180 185 190
55
```

Lys Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser Val Leu Tyr 195 200 205 Asn Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe Gly Gly Lys 210 220 Ala Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val Asn Gly Ile 225 230 235 240 Arg His Ile Gly Leu Ala Ala Lys Gln Leu Glu Gly Ser Gly Gly 245 250 255 10 Gly Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser 260 265 270 Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly 275 280 285 15 Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg 290 295 300 Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro 305 310 315 20 Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys 325 330 335Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg 340 345 350 25 Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile 355 360 365 Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn 370 380 30 Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly 385 390 400 Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu 405 410 415 35 Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His 420 425 430 Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg 435 440 445 40 Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr 450 455 460 Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala 465 470 475 480 45 Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe 485 490 495 Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn 500 505 Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp 515 520 525 50 Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn 530 540 Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr 545 550 555 560 55

Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe 565 570 575 Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp 580 585 590 Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr 595 600 605 Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp 610 620 10 Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn 625 630 635 640 Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly 645 650 655 15 Thr Ala Ala Leu Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn Asp Asn 660 665 670 Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly 675 680 685 20 Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys Ala Met 690 695 700 Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp Thr Lys 705 710 720 25 Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser Gly Thr  $725 \hspace{1cm} 730 \hspace{1cm} 735$ Gly Gly Leu Ile Lys Lys Gly Gly Ser Gln Leu Gln Leu His Gly Asn 740 745 750 30 Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly Ser Leu Val Leu 755 760 765 Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly Ala Leu 770 780 35 Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp Gly Ile 785 790 795 800 Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His 805 810 815 40 Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg 820 825 830 Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys 835 840 845 45 Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr 850 860 Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr 865 870 875 880 Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala Ser Leu 885 890 895 50 Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu Ser Tyr 900 905 910Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala Ala His 915 920 925 55

		Ser	930	Pro	АІа	GIY	Leu	935	His	АІа	vaı	GIU	940	Gly	Gly	Ser	Asn
5		Leu 945	Glu	Asn	Leu	Met	Va1 950	Glu	Leu	Asp	Аlа	Ser 955	Glu	Ser	Ser	Ala	Thr 960
		Pro	Glu	Thr	٧al	G1u 965	Thr	Ala	Ala	Ala	Asp 970	Arg	Thr	Asp	Met	Pro 975	Gly
10	)	Ile	Arg	Pro	Tyr 980	Gly	Ala	Thr	Phe	Arg 985	Ala	Ala	Ala	Аlа	va1 990	Gln	нis
		Ala	Asn	Ala 995	Ala	Asp	Gly	٧a٦	Arg 1000		Phe	Asn	ser	Leu 1009		Ala	Thr
15	5	val	Tyr 1010		Asp	Ser	Thr	Ala 1015		His	Ala	Asp	Met 1020		Gly	Arg	Arg
		Leu 1025		Аlа	val	Ser	Asp 1030	Gly )	Leu	Asp	His	Asn 1035		Thr	Gly	Leu	Arg 1040
20	)	۷al	Ile	Ala	Gln	Thr 1045		Gln	Asp	Gly	Gly 1050		Trp	Glu	Gln	Gly 105	
		٧a٦	Glu	Gly	Lys 1060		Arg	Gly	Ser	Thr 1069		Thr	val	Gly	Ile 1070		Ala
25	5	Lys	⊤hr	Gly 1075		Asn	Thr	Thr	Ala 1080		ΑΊа	⊤hr	Leu	Gly 1085		Glу	Arg
		Ser	Thr 1090		Ser	Glu	Asn	Ser 1099	_	Asn	Ala	Lys	Thr 1100		Ser	Ile	Ser
30	7	Leu 1105		Ala	Gly	Ile	Arg 111(	His )	Asp	Ala	Gly	Asp 1115		Gly	Tyr	Leu	Lys 1120
	,	Gly	Leu	Phe	Ser	Tyr 1125		Arg	Tyr	Lys	Asn 1130		Ile	Ser	Arg	Ser 113	
0.0	_	Gly	Ala	Asp	Glu 1140		Ala	Glu	Gly	Ser 1145		Asn	Gly	Thr	Leu 115(		Gln
35	)	Leu	Gly	Ala 115		Gly	Gly	۷al	Asn 1160		Pro	Phe	Ala	Ala 1165		Gly	Asp
		Leu	Thr 1 <b>17</b> (		Glu	Gly	Gly	Leu 1179		Tyr	Asp	Leu	Leu 1180		Gln	Asp	Ala
40	)	Phe 1185		Glu	Lys	Gly	ser 1190	Ala )	Leu	Gly	тгр	ser 1195		Asn	ser	Leu	⊤hr 1200
		Glu	Gly	Thr	Leu	Val 1205		Leu	Ala	Gly	Leu 1210		Leu	Ser	Gln	Pro 1215	
45	5	Ser	Asp	Lys	Ala 1220		Leu	Phe	Ala	Thr 1225		Gly	val	Glu	Arg 1230		Leu
		Asn	Gly	Arg 1235		Tyr	Thr	Val	Thr 1240		Gly	Phe	Thr	Gly 1249		Thr	Ala
50	)	Ala	Thr 1250		Lys	Thr	Gly	Ala 1255		Asn	Met	Pro	ніs 1260		Arg	Leu	val
		Ala 1265		Leu	Gly	Ala	Asp 1270	val	Glu	Phe	Gly	Asn 1275		Trp	Asn	Gly	Leu 1280
55	5	Ala	Arg	Tyr	Ser	Tyr 1285		Gly	Ser	Lys	Gln 1290		Gly	Asn	His	Ser 1295	

Arg Val Gly Val Gly Tyr Arg Phe Leu Glu His His His His His 1300 1305 1310

```
<210>
                                                                                       114
                                           <211>
                                                                                      2028
                                            <212>
                                                                                     DNA
                                                                                      Artificial Sequence
                                           <213>
                                            <220>
10
                                           <223>
                                                                                     deltaG741-ORF46.1
                                         atggtcgccg ccgacatcgg tgcggggctt gccgatgcac taaccgcacc gctcgaccat aaagacaaag gtttgcagtc tttgacgctg gatcagtccg tcaggaaaaa cgagaaactg aagctggcgg cacaaggtgc ggaaaaaact tatggaaacg gtgacagcct caatacgggc aaattgaaga acgacaaggt cagccgtttc gactttatcc gccaaatcga agtggacggg cagctcatta ccttggagag tggagagttc caagtataca aacaaagcca ttccgcctta accgcctttc agaccgagca aatacaagat tcggagcatt ccgggaagat ggttgcgaaa
                                                                                                                                                                                                                                                                                                                                                                                                                            60
                                                                                                                                                                                                                                                                                                                                                                                                                             120
                                                                                                                                                                                                                                                                                                                                                                                                                            180
15
                                                                                                                                                                                                                                                                                                                                                                                                                              240
                                                                                                                                                                                                                                                                                                                                                                                                                             300
                                                                                                                                                                                                                                                                                                                                                                                                                             360
                                         accgcctttc agaccgagca aatacaagat tcggagcatt ccgggaagat ggttgcgaaa cgccagttca gaatcggcga catagcgggc gaacatacat cttttgacaa gcttcccgaa ggcggcgaggg cgacatatcg cgggacggcg ttcggttcag acgatgccgg cggaaaactg acctacacca tagatttcgc cgccaagcag ggaaacggca aaatcgaaca tttgaaatcg ccagaactca atgtcgacct ggccgccgc gatatcaagc cggatggaaa acgccatgcc gtcatcagcg gttccgtct ttacaaccaa gccgagaaag gcagttactc cctcggtatc tttggcggaa aagcccagga agttgccggc agcgcggaag tgaaaaccgt aaacggcata cgccatatcg gccttgccgc caagcaactc gacggtggcg gaggcactgg atcctcagat ttggcaaacg attctttat ccggcaggt ctcgaccgtc agcattcga acccgacggg aaataccacc tattcggcag cagggggaa cttgccggc gcagcggcca tatcggatggaaaatac aaagccatca gttgggcaac ctgatgatc aacaggcggc cattaaagga aatatcggct acattotcg cttttccgat cacgggcacg aagtccattc ccccttcgac
                                                                                                                                                                                                                                                                                                                                                                                                                             420
                                                                                                                                                                                                                                                                                                                                                                                                                            480
                                                                                                                                                                                                                                                                                                                                                                                                                             540
                                                                                                                                                                                                                                                                                                                                                                                                                            600
20
                                                                                                                                                                                                                                                                                                                                                                                                                            660
                                                                                                                                                                                                                                                                                                                                                                                                                              720
                                                                                                                                                                                                                                                                                                                                                                                                                            780
                                                                                                                                                                                                                                                                                                                                                                                                                            840
                                                                                                                                                                                                                                                                                                                                                                                                                             900
                                                                                                                                                                                                                                                                                                                                                                                                                            960
25
                                         adatated adagecated gittgggedate tigatgatte adaggegge cattadayga anatategget acattgteeg titteegat caegggeacg angecatte eccetteegae anaceatgeet caeatteega tittgatgan geoggeagte eegggeegge attageett tacegeacte attgggaegg atacgaacae cateeegge aeggetatga egggeeacag geoggegget ateeeggee caeaaggeegg agggatatat acagetaega cataanaagge gitgeecaaa atateegget caacetgaee gacaacegea geaceggaea aeggeettgee gacaceeggat acageceegg tagtatgetg aeggeaaggag taggeegge atteaaegge gecaceeggat acageceega getggaeagga egggaeaggat eeggeeggatatataegge geeggaeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataeggataegga
                                                                                                                                                                                                                                                                                                                                                                                                                             1020
                                                                                                                                                                                                                                                                                                                                                                                                                            1080
                                                                                                                                                                                                                                                                                                                                                                                                                            1140
                                                                                                                                                                                                                                                                                                                                                                                                                            1200
                                                                                                                                                                                                                                                                                                                                                                                                                            1260
                                                                                                                                                                                                                                                                                                                                                                                                                            1320
30
                                                                                                                                                                                                                                                                                                                                                                                                                            1380
                                         gccacccgat acagccccga gctggacaga tcgggcaatg ccgccgaagc cttcaacggc actgcagata tcgttaaaaa catcatcggc gcggcaggag aaattgtcgg cgcaggcgat gccgtgcagg gcataagcga aggctcaaac attgctgtca tgcacggctt gggtctgctt tccaccgaaa acaagatggc gcgatcaac gatttggcag atatggcga actcaaagac tatgccgcag cagccatccg cgattgggca gtccaaaacc ccaatgccgc acaaggcata gaagccgtca gcaatactt tatggcagcc atcccatca aagggattgg agctgttcgg ggaaaatacg gcttgggcgg catcacggca catcctatca agcggtcgca gatggcgcg atcgcattgc cgaaagggaa atccgccgtc agcgacaatt ttgccgatgc ggcatacgcc aaatacccgt ccccttacca ttcccgaaat atccgttcaa acttggagca gcgttacggc aacaagaaaaca tcacctcct aaccgtgccg ccgtcaaacg gcaaaaatgt caaactggca gaccaacgcc acccgaagac agcggtaccg tttgacggta aagggttcc gaattttgag aagacacgtga aatatgatac gctcgagcac caccaccac accactga
                                                                                                                                                                                                                                                                                                                                                                                                                            1440
                                                                                                                                                                                                                                                                                                                                                                                                                             1500
                                                                                                                                                                                                                                                                                                                                                                                                                            1560
                                                                                                                                                                                                                                                                                                                                                                                                                            1620
                                                                                                                                                                                                                                                                                                                                                                                                                            1680
                                                                                                                                                                                                                                                                                                                                                                                                                            1740
35
                                                                                                                                                                                                                                                                                                                                                                                                                             1800
                                                                                                                                                                                                                                                                                                                                                                                                                            1860
                                                                                                                                                                                                                                                                                                                                                                                                                            1920
                                                                                                                                                                                                                                                                                                                                                                                                                            1980
                                                                                                                                                                                                                                                                                                                                                                                                                            2028
                                           aagcacgtga aatatgatac gctcgagcac caccaccacc accactga
40
                                           <210>
                                                                                       115
                                           <211>
                                                                                       675
                                            <212>
                                            <213>
                                                                                     Artificial Sequence
                                           <220>
45
                                           <223>
                                                                                      deltaG741-ORF46.1
                                          Met Val Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala
1 5 10 15
50
                                          Pro Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln 20 25 30
                                          Ser Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln Gly Ala Glu
35 40 45
                                          Lys Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys Leu Lys Asn 50 \hspace{1cm} 55
```

Asp Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu Val Asp Gly 65 75 80 Gln Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr Lys Gln Ser 85 90 95 His Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln Asp Ser Glu
100 105 110 His Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile Gly Asp Ile 115 120 125 10 Ala Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala 130 140 Thr Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu 145 150 155 160 15 Thr Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Ile Glu 165 170 175 His Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala Asp Ile 180 185 190 20 Lys Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser Val Leu Tyr 195 200 205Asn Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe Gly Gly Lys 210 215 220 25 Ala Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val Asn Gly Ile 225 230 240 Arg His Ile Gly Leu Ala Ala Lys Gln Leu Asp Gly Gly Gly Thr 245 250 255 30 Gly Ser Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp 260 265 270 Arg Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg 275 280 285 35 Gly Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln 290 295 300 Ser His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly 305 310 315 320 305 40 Asn Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu Val His 325 330 335Ser Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly 340 345 35045 Ser Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr 355 360 365Glu His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Gly Tyr 370 375 380 Pro Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly 385 395 400 50 Val Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly 405 410 415 Gln Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln
420 425 430 55

	Gly Val	Gly Asp 435	Gly	Phe	Lys	Arg 440	Ala	Thr	Arg	туг	Ser 445	Pro	Glu	Leu
5	Asp Arg 450	Ser Gly	' Asn	Аlа	Ala 455	Glu	Ala	Phe	Asn	G]y 460	Thr	Ala	Asp	Ile
	Val Lys 465	Asn Ile	: Ile	Gly 470	Ala	Ala	Gly	Glu	11e 475	٧al	Gly	Ala	Gly	Asp 480
10	Ala Val	Gln Gly	11e 485	Ser	Glu	Gly	ser	Asn 490	Ile	Ala	val	Met	ніs 495	Gly
	Leu Gly	Leu Lei 500		Thr	Glu	Asn	Lys 505	Met	Ala	Arg	Ile	Asn 510	Asp	Leu
15	Ala Asp	Met Ala 515	Gln	Leu	Lys	Asp 520	Tyr	Ala	Ala	Ala	Ala 525	Ile	Arg	Asp
	Trp Ala 530	Val Glr	Asn	Pro	Asn 535	Ala	Ala	Gln	Gly	11e 540	Glu	Ala	Val	Ser
20	Asn Ile 545	Phe Met	: Ala	Ala 550	Ile	Pro	Ile	Lys	G]y 555	Ile	Gly	Ala	val	Arg 560
	Gly Lys	Tyr Gly	Leu 565	Gly	Gly	ıle	Thr	Ala 570	нis	Pro	Ile	Lys	Arg 575	ser
25	Gln Met	Gly Ala 580		A∃a	Leu	Pro	Lys 585	Gly	Lys	Ser	Ala	va1 590	Ser	Asp
	Asn Phe	Ala Asp 595	Ala	Ala	Tyr	Ala 600	Lys	Tyr	Pro	Ser	Pro 605	Tyr	His	Ser
30	Arg Asn 610	Ile Arg	, Ser	Asn	Leu 615	Glu	Gln	Arg	Tyr	Gly 620	Lys	Glu	Asn	Ile
	Thr Ser 625	Ser Thi	· val	Pro 630	Pro	Ser	Asn	Gly	Lys 635	Asn	٧al	Lys	Leu	Ala 640
<i>35</i>	Asp Gln	Arg His	645	Lys	Thr	Gly	val	Pro 650	Phe	Asp	Gly	Lys	G]y 655	Phe
	Pro Asn	Phe Glu 660		His	val	Lys	Tyr 665	Asp	Thr	Leu	Glu	Нis 670	His	His
40	His His	His 675												
	<210> <211> <212> <213>	116 249 PRT Artific	ial:	Sequ	ence									
45	<220> <223>	Novel p	rote	in										
	<400> Met Lys 1	116 Lys Tyl	Leu 5	Phe	Arg	Ala	Аlа	Leu 10	Tyr	Glу	Ile	Ala	Ala 15	Ala
50	Ile Fen	Ala Ala 20	ı Ala	Ile	Pro	Аlа	Gly 25	Asn	Asp	Ala	Thr	Thr 30	Lys	Pro
	Asp Leu	Tyr Tyi 35	Leu	Lys	Asn	Glu 40	Gln	Ala	Ile	Asp	Ser 45	Leu	Lys	Leu
55	Leu Pro 50	Pro Pro	) Pro	Glu	Va 1 55	Gly	Ser	Ile	Gln	Phe 60	Leu	Asn	Asp	Gln

```
Ala Met Tyr Glu Lys Gly Arg Met Leu Arg Asn Thr Glu Arg Gly Lys
65 75 80
          Gln Ala Gln Ala Asp Ala Asp Leu Ala Ala Gly Gly Val Ala Thr Ala
85 90 95
          Phe Ser Gly Ala Phe Gly Tyr Pro Ile Thr Glu Lys Asp Ser Pro Glu 100 105 110
          Leu Tyr Lys Leu Leu Thr Asn Met Ile Glu Asp Ala Gly Asp Leu Ala
115 120 125
10
          Thr Arg Ser Ala Lys Glu His Tyr Met Arg Ile Arg Pro Phe Ala Phe 130 140
          Tyr Gly Thr Glu Thr Cys Asn Thr Lys Asp Gln Lys Lys Leu Ser Thr
145 150 155
15
          Asn Gly Ser Tyr Pro Ser Gly His Thr Ser Ile Gly Trp Ala Thr Ala
165 170 175
          Leu Val Leu Ala Glu Val Asn Pro Ala Asn Gln Asp Ala Ile Leu Glu
180 185 190
20
          Arg Gly Tyr Gln Leu Gly Gln Ser Arg Val Ile Cys Gly Tyr His Trp
195 200 205
          Gln Ser Asp Val Asp Ala Ala Arg Ile Val Gly Ser Ala Ala Val Ala
210 220
25
          Thr Leu His Ser Asp Pro Ala Phe Gln Ala Gln Leu Ala Lys Ala Lys 225 230 235 240
          Gln Glu Phe Ala Gln Lys Ser Gln Lys
245
30
          <210>
                    117
          <211>
                    66
          <212>
                    DNA
                    Artificial Sequence
          <220>
35
          <223>
                    L1 linker
          <220>
          <221>
<222>
                    13
                    \overline{A}, T/U, G or C
           <223>
40
          <400>
          tatgaartay ytnttymgcg ccgccctgta cggcatcgcc gccgccatcc tcgccgccgc
                                                                                           60
                                                                                           66
          gatccc
          <210>
                    118
          <211>
                    69
45
          <212>
                    DNA
                    Artificial Sequence
          <213>
          <220>
<223>
                    S1 linker
50
          <220>
          <221>
                    25, 28
A, T/U, G or C
           <222>
           <223>
          <400>
55
          tatgaaaaaa tacctattcc grgcngcnyt rtayggsatc gccgccgcca tcctcgccgc
                                                                                           60
                                                                                           69
          cgcgatccc
```

```
<210>
                119
       <211>
                27
        <212>
                DNA
        <213>
                Artificial Sequence
       <220>
<223>
                9L1-a
       <400>
                119
       atgaagaagt accttttcag cgccgcc
                                                                                   27
10
       <210>
                120
       <211>
                27
       <212>
                DNA
        <213>
                Artificial Sequence
       <220>
15
       <223>
                9L1-e
       <400>
                120
                                                                                   27
       atgaaaaaat actttttccg cgccgcc
                121
27
       <210>
20
       <211>
       <212>
                DNA
                Artificial Sequence
       <213>
       <220>
       <223>
                9L1-d
25
        <400>
                121
       atgaaaaaat actttttccg cgccgcc
                                                                                   27
       <210>
                122
       <211>
                60
        <212>
30
                DNA
        <213>
                Artificial Sequence
       <220>
       <223>
                9L1-f
       <400>
35
       atgaaaaat atctctttag cgccgccctg tacggcatcg ccgccgccat cctcgccgcc
                                                                                   60
       <210>
                123
        <211>
                60
        <212>
                DNA
       <213>
                Artificial Sequence
40
       <220>
<223>
                919sp
       <400>
                123
       atgaaaaaat acctattccg cgccgccctg tacggcatcg ccgccgccat cctcgccgcc
45
       <210>
                124
       <211>
<212>
                PRT
       <213>
                Artificial Sequence
50
       <220>
       <223>
                9L1a
        <400>
                124
       Met Lys Lys Tyr Leu Phe Ser Ala Ala
55
       <210>
                125
```

```
<211>
                  PRT
         <212>
                  Artificial Sequence
         <220>
         <223>
                  9L1e
         <400>
                  125
         Met Lys Lys Tyr Phe Phe Arg Ala Ala
1 5
10
         <210>
                  126
9
         <211>
<212>
                  PRT
                  Artificial Sequence
         <213>
15
         <220>
                  9L1d
         <223>
         <400>
         Met Lys Lys Tyr Phe Phe Arg Ala Ala
1 5
20
         <210>
                  127
20
         <211>
<212>
                  PRT
                  Artificial Sequence
         <213>
         <220>
25
                  9L1f
         <223>
         <400>
         Met Lys Lys Tyr Leu Phe Ser Ala Ala Leu Tyr Gly Ile Ala Ala Ala 1 5 10 15
30
         Ile Leu Ala Ala
20
         <210>
                  128
         <211>
                  20
         <212>
         <213>
                  Artificial Sequence
35
         <220>
         <223>
                  9L1sp
         <400>
         Met Lys Lys Tyr Leu Phe Arg Ala Ala Leu Tyr Gly Ile Ala Ala 1 10 15
40
         Ile Leu Ala Ala
         <210>
                  129
45
         <211>
                  42
         <212>
                  DNA
                  Artificial Sequence
         <213>
         <220>
         <223>
                  9s1-e
50
         <400>
                  129
                                                                                      42
         atgaaaaaat acctattcat cgccgccgcc atcctcgccg cc
         <210>
                  130
         <211>
                  60
         <212>
                  DNA
55
                  Artificial Sequence
         <213>
```

```
<220>
                   9s1-c
          <223>
          <400>
                   130
5
          atgaaaaaat acctattccg agctgcccaa tacggcatcg ccgccgccat cctcgccgcc
                                                                                     60
          <210>
                   131
          <211><212>
                   DNA
                   Artificial Sequence
          <213>
10
          <220>
                   9s1-b
          <223>
          <400>
                   131
          atgaaaaaat acctattccg ggccgcccaa tacggcatcg ccgccgccat cctcgccgcc
                                                                                     60
15
          <210>
                   132
          <211>
                   60
          <212>
                   DNA
                   Artificial Sequence
          <213>
          <220>
20
                   9s1-i
          <223>
          <400>
                   132
          atgaaaaaat acctattccg ggcggctttg tacgggatcg ccgccgccat cctcgccgcc
          <210>
25
          <211>
                   14
          <212>
                   PRT
                   Artificial Sequence
          <213>
          <220>
          <223>
                   9s1e
30
          <400>
          Met Lys Lys Tyr Leu Phe Ile Ala Ala Ile Leu Ala Ala 1
          <210>
          <211>
                   20
35
          <212>
                   PRT
          <213>
                   Artificial Sequence
          <220>
                   951c
          <223>
40
          <400>
          Met Lys Lys Tyr Leu Phe Arg Ala Ala Gln Tyr Gly Ile Ala Ala Ala 1 5 10 15
          Ile Leu Ala Ala
20
45
          <210>
                   135
          <211>
                   20
          <212>
                   PRT
                   Artificial Sequence
          <213>
          <220>
50
          <223>
                   9s1b
          <400>
          Met Lys Lys Tyr Leu Phe Arg Ala Ala Gln Tyr Gly Ile Ala Ala Ala 1 5 10 15
55
          Ile Leu Ala Ala
```

```
<210>
                     136
            <211>
                     20
            <212>
                     Artificial Sequence
            <220>
           <223>
                     951i
            <400>
           Met Lys Lys Tyr Leu Phe Arg Ala Ala Leu Tyr Gly Ile Ala Ala Ala
1 5 10 15
10
           Ile Leu Ala Ala
20
           <210>
15
            <211>
                     467
            <212>
                     PRT
            <213>
                     Artificial Sequence
           <220>
                     730
           <223>
20
           Val Lys Pro Leu Arg Arg Leu Thr Asn Leu Leu Ala Ala Cys Ala Val
1 5 10 15
           Ala Ala Ala Leu Ile Gln Pro Ala Leu Ala Ala Asp Leu Ala Gln
20 25 30
25
           Asp Pro Phe Ile Thr Asp Asn Ala Gln Arg Gln His Tyr Glu Pro Gly 35 40 45
           Gly Lys Tyr His Leu Phe Gly Asp Pro Arg Gly Ser Val Ser Asp Arg 50 60
30
           Thr Gly Lys Ile Asn Val Ile Gln Asp Tyr Thr His Gln Met Gly Asn 65 70 75 80
           Leu Leu Ile Gln Gln Ala Asn Ile Asn Gly Thr Ile Gly Tyr His Thr 85 90 95
35
           Arg Phe Ser Gly His Gly His Glu Glu His Ala Pro Phe Asp Asn His 100 105 110
           Ala Ala Asp Ser Ala Ser Glu Glu Lys Gly Asn Val Asp Glu Gly Phe 115 120 125
40
           Thr Val Tyr Arg Leu Asn Trp Glu Gly His Glu His His Pro Ala Asp
130 135 140
           Ala Tyr Asp Gly Pro Lys Gly Gly Asn Tyr Pro Lys Pro Thr Gly Ala
145 150 155 160
45
           Arg Asp Glu Tyr Thr Tyr His Val Asn Gly Thr Ala Arg Ser Ile Lys
165 170 175
           Leu Asn Pro Thr Asp Thr Arg Ser Ile Arg Gln Arg Ile Ser Asp Asn 180 \hspace{1cm} 185 \hspace{1cm} 190
50
           Tyr Ser Asn Leu Gly Ser Asn Phe Ser Asp Arg Ala Asp Glu Ala Asn
195 200 205
           Arg Lys Met Phe Glu His Asn Ala Lys Leu Asp Arg Trp Gly Asn Ser 210 215 220
           Met Glu Phe Ile Asn Gly Val Ala Ala Gly Ala Leu Asn Pro Phe Ile
225 230 235
55
```

Ser Ala Gly Glu Ala Leu Gly Ile Gly Asp Ile Leu Tyr Gly Thr Arg 245 250 255 Tyr Ala Ile Asp Lys Ala Ala Met Arg Asn Ile Ala Pro Leu Pro Ala 260 270 Glu Gly Lys Phe Ala Val Ile Gly Gly Leu Gly Ser Val Ala Gly Phe 275 280 285 Glu Lys Asn Thr Arg Glu Ala Val Asp Arg Trp Ile Gln Glu Asn Pro 290 295 300 10 Asn Ala Ala Glu Thr Val Glu Ala Val Phe Asn Val Ala Ala Ala 305 310 315 320 Lys Val Ala Lys Leu Ala Lys Ala Ala Lys Pro Gly Lys Ala Ala Val 325 330 335 15 Ser Gly Asp Phe Ala Asp Ser Tyr Lys Lys Leu Ala Leu Ser Asp 340 345 Leu Ala Leu Ser Asp Ser Ala Arg Gln Leu Tyr Gln Asn Ala Lys Tyr Arg Glu Ala Leu Asp 355 360 36520 Ile His Tyr Glu Asp Leu Ile Arg Arg Lys Thr Asp Gly Ser Ser Lys 370 375 380Phe Ile Asn Gly Arg Glu Ile Asp Ala Val Thr Asn Asp Ala Leu Ile 385 390 395 400 25 Gln Ala Lys Arg Thr Ile Ser Ala Ile Asp Lys Pro Lys Asn Phe Leu 405 410 415Asn Gln Lys Asn Arg Lys Gln Ile Lys Ala Thr Ile Glu Ala Ala Asn 420 425 430 30 Gln Gln Gly Lys Arg Ala Glu Phe Trp Phe Lys Tyr Gly Val His Ser 435 440 445 Gln Val Lys Ser Tyr Ile Glu Ser Lys Gly Gly Ile Val Lys Thr Gly
450 460 35 Leu Gly Asp 465 <210> 138 <211> 377 <212> 40 PRT Artificial Sequence <213> <220> 730-c1 <223> <400> 45 Met Ala Asp Leu Ala Gln Asp Pro Phe Ile Thr Asp Asn Ala Gln Arg
1 10 15 Gln His Tyr Glu Pro Gly Gly Lys Tyr His Leu Phe Gly Asp Pro Arg 20 25 30 50 Gly Ser Val Ser Asp Arg Thr Gly Lys Ile Asn Val Ile Gln Asp Tyr 35 40 Thr His Gln Met Gly Asn Leu Leu Ile Gln Gln Ala Asn Ile Asn Gly 50 60 55 Thr Ile Gly Tyr His Thr Arg Phe Ser Gly His Gly His Glu Glu His 65 70 75 80

```
Ala Pro Phe Asp Asn His Ala Ala Asp Ser Ala Ser Glu Glu Lys Gly
85 90 95
            Asn Val Asp Glu Gly Phe Thr Val Tyr Arg Leu Asn Trp Glu Gly His
100 105 110
            Glu His His Pro Ala Asp Ala Tyr Asp Gly Pro Lys Gly Gly Asn Tyr
115 120 125
             Pro Lys Pro Thr Gly Ala Arg Asp Glu Tyr Thr Tyr His Val Asn Gly 130 140
10
            Thr Ala Arg Ser Ile Lys Leu Asn Pro Thr Asp Thr Arg Ser Ile Arg
145 150 155 160
            Gln Arg Ile Ser Asp Asn Tyr Ser Asn Leu Gly Ser Asn Phe Ser Asp 165 170 175
15
            Arg Ala Asp Glu Ala Asn Arg Lys Met Phe Glu His Asn Ala Lys Leu
180 185 190
            Asp Arg Trp Gly Asn Ser Met Glu Phe Ile Asn Gly Val Ala Ala Gly 195 200 205
20
            Ala Leu Asn Pro Phe Ile Ser Ala Gly Glu Ala Leu Gly Ile Gly Asp 210 220
             Ile Leu Tyr Gly Thr Arg Tyr Ala Ile Asp Lys Ala Ala Met Arg Asn
225 230 240
25
            Ile Ala Pro Leu Pro Ala Glu Gly Lys Phe Ala Val Ile Gly Gly Leu
245 250 255
            Gly Ser Val Ala Gly Phe Glu Lys Asn Thr Arg Glu Ala Val Asp Arg
260 265 270
30
            Trp Ile Gln Glu Asn Pro Asn Ala Ala Glu Thr Val Glu Ala Val Phe 275 280 285
            Asn Val Ala Ala Ala Ala Lys Val Ala Lys Leu Ala Lys Ala Ala Lys 290 295 300
35
             Pro Gly Lys Ala Ala Val Ser Gly Asp Phe Ala Asp Ser Tyr Lys Lys 305 310 315 320
            Lys Leu Ala Leu Ser Asp Ser Ala Arg Gln Leu Tyr Gln Asn Ala Lys 325 330 335
40
            Tyr Arg Glu Ala Leu Asp Ile His Tyr Glu Asp Leu Ile Arg Arg Lys 340 345 350
            Thr Asp Gly Ser Ser Lys Phe Ile Asn Gly Arg Glu Ile Asp Ala Val 355 360 365
45
            Thr Asn Asp Ala Leu Ile Gln Ala Arg
370 375
                      139
353
             <210>
             <211>
             <212>
                      PRT
50
                      Artificial Sequence
             <220>
             <223>
                      730-C2
55
            Met Ala Asp Leu Ala Gln Asp Pro Phe Ile Thr Asp Asn Ala Gln Arg
1 10 15
```

Gln His Tyr Glu Pro Gly Gly Lys Tyr His Leu Phe Gly Asp Pro Arg Gly Ser Val Ser Asp Arg Thr Gly Lys Ile Asn Val Ile Gln Asp Tyr 35 40 Thr His Gln Met Gly Asn Leu Leu Ile Gln Gln Ala Asn Ile Asn Gly 50 60 Thr Ile Gly Tyr His Thr Arg Phe Ser Gly His Gly His Glu Glu His 65 70 75 80 10 Ala Pro Phe Asp Asn His Ala Ala Asp Ser Ala Ser Glu Glu Lys Gly 85 90 95 Asn Val Asp Glu Gly Phe Thr Val Tyr Arg Leu Asn Trp Glu Gly His 100 105 110 15 Glu His His Pro Ala Asp Ala Tyr Asp Gly Pro Lys Gly Gly Asn Tyr 115 120 125 Pro Lys Pro Thr Gly Ala Arg Asp Glu Tyr Thr Tyr His Val Asn Gly 130 135 140 20 Thr Ala Arg Ser Ile Lys Leu Asn Pro Thr Asp Thr Arg Ser Ile Arg 145 150 155 160 Gln Arg Ile Ser Asp Asn Tyr Ser Asn Leu Gly Ser Asn Phe Ser Asp 165 170 25 Arg Ala Asp Glu Ala Asn Arg Lys Met Phe Glu His Asn Ala Lys Leu 180 185 190 Asp Arg Trp Gly Asn Ser Met Glu Phe Ile Asn Gly Val Ala Ala Gly 195 200 205 30 Ala Leu Asn Pro Phe Ile Ser Ala Gly Glu Ala Leu Gly Ile Gly Asp 210 220 Ile Leu Tyr Gly Thr Arg Tyr Ala Ile Asp Lys Ala Ala Met Arg Asn 225 230 235 240 35 Ile Ala Pro Leu Pro Ala Glu Gly Lys Phe Ala Val Ile Gly Gly Leu 245 250 255 Gly Ser Val Ala Gly Phe Glu Lys Asn Thr Arg Glu Ala Val Asp Arg 260 265 270 40 Trp Ile Gln Glu Asn Pro Asn Ala Ala Glu Thr Val Glu Ala Val Phe 275 280 285 Asn Val Ala Ala Ala Ala Lys Val Ala Lys Leu Ala Lys Ala Ala Lys 290 295 300 45 Pro Gly Lys Ala Ala Val Ser Gly Asp Phe Ala Asp Ser Tyr Lys Lys 305 310 315 Lys Leu Ala Leu Ser Asp Ser Ala Arg Gln Leu Tyr Gln Asn Ala Lys 325 330 335 50 Tyr Arg Glu Ala Leu Gly Lys Val Arg Ile Ser Gly Glu Ile Leu Leu 340 345 350 Gly 55 <210> 140 <211> 2019

```
<212>
                       <213>
                                            Artificial Sequence
                       <220>
                      <223>
                                            ORF46.1-741
                      <400>
                      atgtcagatt tggcaaacga ttcttttatc cggcaggttc tcgaccgtca gcatttcgaa
                                                                                                                                                                                                                60
                      cccgacggga aataccacct attcggcagc aggggggaac ttgccgagcg cagcggccat
                                                                                                                                                                                                                 120
                      atcggattgg gaaaaataca aagccatcag ttgggcaacc tgatgattca acaggcggcc attaaaggaa atatcggcta cattgtccgc ttttccgatc acgggcacga agtccattcc ccttcgaca accatgcctc acattccgat tctgatgaag ccggtagtcc cgttgacgga ttagccctt accgcatca ttgggatga tagacgacacc atccgccga cggctatgac
                                                                                                                                                                                                                180
                                                                                                                                                                                                                 240
10
                                                                                                                                                                                                                 300
                                                                                                                                                                                                                 360
                     tttagccttt accgcatcca ttgggacgga tacgaacac atcccgccga cggctatgac gggccacagg gcggcggcta tcccgctcc aaaggcgcga gggatatata cagctacgac ataaaaggcg ttgcccaaaa tatccgcctc aacctgaccg acaaccgcag caccggacaa cggcttgccg accgtttcca caatgccggt agtatgctga cgcaaggagt aggcgacggc ttcaaacggc ccacccgata cagccccgag ctggacagat cggcaaggc cgcaggaga ctgcaggaga ctgcaggaga ctgcaggaga cggtaaaac atcatcggcg cggcaggaga aattgtcggc gcaggcgatg ccgtgcaggg cataagcgaa ggctcaaaca ttgctgtcat gcacggcttg ggtctgcttt ccaccgaaaa caagatggcg cgcatcaaca ttgctgtcat gcacggcttg aggcctgttggaa aatgccgcag agccatccgc gattgggcag tccaaaaccc caatgccgca caaggcatag aagccgtcag caatatctt atggcagca tcccatcaa agggattgga gctgtcggg gaaaatacgg cttgggcggca atcacggca acctatcaa gcggtcggag tcgcattggcag atcggcgcaa tcgcattgcc gaaagggaaa tccgcgcaa accatcaca accgccaattt tgccgatgcg gcatacgca aatacccgtc cccttaccaa tccccgaaata tccgttcaaa cttggagcag cgttacggca aagaaaacat cacctcctca accggccc cgtcaaacgg caaaaatgtc
                                                                                                                                                                                                                 420
                                                                                                                                                                                                                 480
                                                                                                                                                                                                                 540
                                                                                                                                                                                                                600
15
                                                                                                                                                                                                                660
                                                                                                                                                                                                                720
                                                                                                                                                                                                                780
                                                                                                                                                                                                                840
                                                                                                                                                                                                                 900
                                                                                                                                                                                                                960
20
                                                                                                                                                                                                                1020
                                                                                                                                                                                                                1080
                      čgttacggca aagaaaacat cacctcctca accgigccgc cgicaaacgg caaaaaigtc
                                                                                                                                                                                                                1140
                      aaactggcag accaacgcca cccgaagaca ggcgtaccgt ttgacggtaa agggtttccg aattttgaga agcacgtgaa atatgatacg ggatccggag ggggtggtgt cgccgccgac atcggtgcgg ggcttgccga tgcactaacc gcaccgctcg accataaaga caaaggtttg cagtctttga cgctggatca gtccgtcagg aaaaacgaga aacttatgg gaacggtgac agcctcaata cgggcaaatt gaagaacgac aaggtcagcc gtttcgactt tatccgccaa atcgaagtg acgggcagct cattaccttg
                                                                                                                                                                                                                1200
1260
                                                                                                                                                                                                                 1320
25
                                                                                                                                                                                                                1380
1440
                                                                                                                                                                                                                1500
                      aaggtcagcc gtttcgactt tatccgccaa atcgaagtgg acgggcagct cattaccttg gagagtggag agttccaagt atacaaacaa agccattccg ccttaaccgc ctttcagacc gagcaaatac aagattcgga gcattccggg aagatggttg cgaaacgcca gttcagaatc ggcgacatag cgggcgaaca tacatcttt gacaagcttc ccgaaggcgg cagggcgaca tatcgcggga cggcgttcgg ttcagacgat gccggcggaa aactgaccta caccatagat ttcgccgcca agcagggaaa cggcaaaatc gaacatttga aatcgccaga actcaatgtc gacctggccg ccgccgatat caagccggat ggaaaacgcc atgccgtcat cagcggttcc gtcctttaca accaagccga gaaaggcagt tactccctcg gtatctttgg cggaaaagcc
                                                                                                                                                                                                                1560
                                                                                                                                                                                                                1620
                                                                                                                                                                                                                1680
30
                                                                                                                                                                                                                1740
                                                                                                                                                                                                                 1800
                                                                                                                                                                                                                1860
1920
                      caggaagttg ccggcagcgc ggaagtgaaa accgtaaacg gcatacgcca tatcggcctt gccgccaagc aactcgagca ccaccaccac caccactga
                                                                                                                                                                                                                 1980
                                                                                                                                                                                                                 2019
35
                      <210>
                                            672
                      <211>
                       <212>
                                            PRT
                       <213>
                                            Artificial Sequence
                       <220>
40
                      <223>
                                            ORF46.1-741
                      Met Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg
1 5 10 15
45
                      Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg Gly 20 25 30
                      Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln Ser 35 40 45
50
                      His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly Asn 50 55 60
                      Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu Val His Ser
65 70 75 80
                      Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly Ser
85 90 95
55
```

Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu 100 105 110 His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Gly Tyr Pro 115 120 125 Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val 130 135 140 Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln 145 150 155 10 Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly 165 170 175 Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp 180 185 19015 Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val 195 200 205 Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp Ala 210 215 220 20 Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu 225 230 240 Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala 245 250 255 25 Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ala Ile Arg Asp Trp 260 265 270 Ala Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn 275 280 285 30 Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly 290 295 300 Lys Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln 305 310 315 320 35 Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn 325 330 335 Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg 340 345 350 40 Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr 355 360 365 Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp 370 375 380 Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro 385 390 395 400 45 Asn Phe Glu Lys His Val Lys Tyr Asp Thr Gly Ser Gly Gly Gly 415 Val Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala Pro 420 425 430 50 Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln Ser 435 440 445 Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln Gly Ala Glu Lys 450 460 55

```
Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys Leu Lys Asn Asp 465 470 475
            Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu Val Asp Gly Gln 485 490 495
            Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr Lys Gln Ser His 500 505 510
            Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln Asp Ser Glu His 515 520 525
10
            Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile Gly Asp Ile Ala 530 540
            Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala Thr 545 550 555 560
            Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu Thr 565 570
            Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Ile Glu His 580 585 590
20
            Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala Asp Ile Lys 595 600
            Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser Val Leu Tyr Asn 610 620
            Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe Gly Gly Lys Ala
625 630 635 640
            Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val Asn Gly Ile Arg
645 650
30
            His Ile Gly Leu Ala Ala Lys Gln Leu Glu His His His His His 660 665 670
            <210>
                        142
                        2421
            <211>
35
             <212>
                        DNA
             <213>
                        Artificial Sequence
            <220>
            <223>
                        ORF46.1-961
            <400>
40
            atgtcagatt tggcaaacga ttcttttatc cggcaggttc tcgaccgtca gcatttcgaa cccgacggga aataccacct attcggcagc aggggggaac ttgccgagcg cagcggccat
                                                                                                                60
                                                                                                                120
            atcggattgg gaaaaataca aagccatcag ttgggcaacc tgatgattca acaggcggcc attaaaggaa atatcggcta cattgtccgc ttttccgatc acgggcacga agtccattcc cccttcgaca accatgcctc acattccgat tctgatgaag ccggtagtcc cgttgacgga
                                                                                                                180
                                                                                                                240
                                                                                                                300
            tttagcčiti accgcátcca tigggacóga tacógačacó atóccgócga cógcéatójác
                                                                                                                360
45
            gggccacagg gcggcggcta tcccgctccc aaaggcgcga gggatatata cagctacgac
ataaaaggcg ttgcccaaaa tatccgcctc aacctgaccg acaaccgcag caccggacaa
                                                                                                                420
                                                                                                                480
            cggcttgccg accgtttcca caatgccggt agtatgctga cgcaaggagt aggcgacgga
                                                                                                                540
            ttcaaacgcg ccacccgata cagccccgag ctggacagat cgggcaatgc cgccgaagcc
ttcaacggca ctgcagatat cgttaaaaac atcatcggcg cggcaggaga aattgtcggc
                                                                                                                600
                                                                                                                660
            gcaggcgatg ccgtgcaggg cataagcgaa ggctcaaaca ttgctgtcat gcacggcttg
                                                                                                                720
50
            ggtctgcttt ccaccgaaaa caagatggcg cgcatcaacg atttggcaga tatggcgcaa ctcaaagact atgccgcagc agccatccgc gattgggcag tccaaaaccc caatgccgca caaggcatag aagccgtcag caatatcttt atggcagcca tccccatcaa agggattgga
                                                                                                                780
                                                                                                                840
                                                                                                                900
            gctgttcggg gaaaatacgg cttgggcggc atcacggcac atcctatcaa gcggtcgcag atgggcgcga tcgcattgcc gaaagggaaa tccgccgtca gcgacaattt tgccgatgcg gcatacgcca aatacccgtc cccttaccat tcccgaaata tccgttcaaa cttggagcag
                                                                                                                960
                                                                                                                1020
                                                                                                                1080
55
            cgttacggca aagaaaacat cacctcctca accgtgccgc cgtcaaacgg caaaaatgtc
                                                                                                               1140
            adactggčag accaacgcca cccgaagaca ggcgtaccgt ttgacggtad agggtttccg
                                                                                                                1200
```

1260

```
aattttgaga agcacgtgaa atatgatacg ggatccggag gaggaggagc cacaaacgac
gacgatgtta aaaaagctgc cactgtggcc attgctgctg cctacaacaa tggccaagaa
atcaacggtt tcaaagctgg agagaccatc tacgacattg atgaagacgg cacaattacc
                                                                                                             1320
                                                                                                            1380
            aaaaaagacg caactgcagc cgatgttgaa gccgacgact ttaaaggtct gggtctgaaa
                                                                                                             1440
            aaagtcgtga ctaacctgac caaaaccgtc aatgaaaaca aacaaaacgt cgatgccaaa gtaaaagctg cagaatctga aatagaaaag ttaacaacca agttagcaga cactgatgcc
                                                                                                             1500
                                                                                                             1560
            gctttagcag atactgatgc cgctctggat gcaaccacca acgccttgaa taaattggga gaaaatataa cgacatttgc tgaagagact aagacaaata tcgtaaaaat tgatgaaaaa ttagaagccg tggctgatac cgtcgacaag catgccgaag cattcaacga tatcgccgat
                                                                                                             1620
                                                                                                             1680
                                                                                                             1740
            tcattggatg aaaccaacac taaggcagac gaagccgtca aaaccgccaa tgaagccaaa
                                                                                                             1800
            cagacggccg aagaaaccaa acaaaacgtc gatgccaaag taaaagctgc agaaactgca
gcaggcaaag ccgaagctgc cgctggcaca gctaatactg cagccgacaa ggccgaagct
10
                                                                                                             1860
                                                                                                             1920
            gtcgctgcaa aagttaccga catcaaagct gatatcgcta cgaacaaaga taatattgct
                                                                                                             1980
            aaaaaagcaa acagtgccga cgtgtacacc agagaagagt ctgacagcaa atttgtcaga attgatggtc tgaacgctac taccgaaaaa ttggacacac gcttggcttc tgctgaaaaa
                                                                                                             2040
                                                                                                             2100
            tccattgccg atcacgatac tcgcctgaac ggtttggata aaacagtgtc agacctgcgc
                                                                                                             2160
            aaagaaaccc gccaaggcct tgcagaacaa gccgcgctct ccggtctgtt ccaaccttac
                                                                                                             2220
            aacgtgggtc ggttcaatgt aacggctgca gtcggcggct acaaatccga atcggcagtc
                                                                                                             2280
            gccatcggta ccggcttccg ctttaccgaa aactttgccg ccaaagcagg cgtggcagtcggcacttcgt ccggttcttc cgcagcctac catgtcggcg tcaattacga gtggctcgag
                                                                                                             2340
                                                                                                            2400
            caccaccacc accaccactg a
                                                                                                            2421
            <210>
                       143
20
            <211>
                       806
            <212>
                       PRT
                       Artificial Sequence
            <213>
            <220>
                       ORF46.1-961
            <223>
            <400>
            Met Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg 1 \hspace{1cm} 5 \hspace{1cm} 10 \hspace{1cm} 15
            Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg Gly 20 25 30
30
            Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln Ser 35 40
            His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ile Lys Gly Asn 50 55
35
            Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu Val His Ser 65 70 75 80
            Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly Ser
85 90 95
40
            Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu 100 105 110
            His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Gly Tyr Pro
115 120
45
            Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val
130 140
            Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln 145 150 155
50
            Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly
165 170 175
            Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp
180 185 190
            Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val
195 200 205
```

	Lys	Asn 210	Ile	Ile	Gly	ΑΊа	Ala 215	Gly	Glu	Ile	٧al	G]y 220	Ala	Gly	Asp	Ala
5	Val 225	Gln	Gly	Ile	Ser	G]u 230	Gly	Ser	Asn	Ile	Ala 235	۷a٦	Met	His	Gly	Leu 240
	Gly	Leu	Leu	Ser	Thr 245	Glu	Asn	Lys	Met	Ala 250	Arg	Ile	Asn	Asp	Leu 255	Ala
10	Asp	Met	Ala	G]n 260	Leu	Lys	Asp	Tyr	Ala 265	Ala	Ala	Ala	Ile	Arg 270	Asp	Trp
	Ala	۷a٦	G]n 275	Asn	Pro	Asn	Ala	Ala 280	Gln	Gly	Ile	Glu	Ala 285	val	Ser	Asn
15	Ile	Phe 290	Met	Ala	Ala	Ile	Pro 295	Ile	Lys	Gly	Ile	Gly 300	Ala	٧a٦	Arg	Gly
	Lys 305	Tyr	GТу	Leu	G1y	Gly 310	Ile	Thr	Ala	His	Pro 315	Ile	Lys	Arg	Ser	G1n 320
20	Met	Gly	Аlа	Ile	Аlа 325	Leu	Pro	Lys	Gly	Lys 330	Ser	Аlа	val	Ser	Asp 335	Asn
	Phe	Ala	Asp	Ala 340	Ala	Tyr	Ala	Lys	Tyr 345	Pro	Ser	Pro	Tyr	нis 350	Ser	Arg
25	Asn	Ile	Arg 355	Ser	Asn	Leu	Glu	G1n 360	Arg	Tyr	Gly	Lys	Glu 365	Asn	Ile	Thr
	Ser	Ser 370	Thr	Val	Pro	Pro	Ser 375	Asn	Gly	Lys	Asn	Val 380	Lys	Leu	Ala	Asp
30	Gln 385	Arg	His	Pro	Lys	Thr 390	Gly	Val	Pro	Phe	Asp 395	Gly	Lys	Gly	Phe	Pro 400
	Asn	Phe	Glu	Lys	His 405	val	Lys	Tyr	Asp	Thr 410	Gly	Ser	Gly	Gly	Gly 415	Gly
35	Ala	Thr	Asn	Asp 420	Asp	Asp	val	Lys	Lys 425	Ala	Ala	Thr	val	А1а 430	Ile	Ala
	Ala	Ala	Tyr 435	Asn	Asn	Gly	Gln	Glu 440	Ile	Asn	Gly	Phe	Lys 445	Ala	Gly	Glu
40	Thr	11e 450	Туг	Asp	Ile	Asp	Glu 455	Asp	Gly	Thr	Ile	Thr 460	Lys	Lys	Asp	Ala
	Thr 465	Ala	Ala	Asp	val	G1u 470	Ala	Asp	Asp	Phe	Lys 475	Gly	Leu	Gly	Leu	Lys 480
45	Lys	val	val	Thr	Asn 485	Leu	Thr	Lys	Thr	va1 490	Asn	Glu	Asn	Lys	Gln 495	Asn
	Val	Asp	Ala	Lys 500	۷a٦	Lys	Ala	Ala	G1u 505	Ser	Glu	Ile	Glu	Lys 510	Leu	Thr
50	Thr	Lys	Leu 515	Ala	Asp	Thr	Asp	А1а 520	Ala	Leu	Ala	Asp	Thr 525	Asp	Ala	Ala
50	Leu	Asp 530	Ala	Thr	Thr	Asn	Ala 535	Leu	Asn	Lys	Leu	Gly 540	Glu	Asn	Ile	Thr
	Thr 545	Phe	Ala	Glu	Glu	Thr 550	Lys	Thr	Asn	ıle	val 555	Lys	Ile	Asp	Glu	Lys 560
55	Leu	Glu	Ala	٧a٦	Ala 565	Asp	Thr	val	Asp	Lys 570	His	Ala	Glu	Ala	Phe 575	Asn

```
Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu Ala 580 590
         Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys Gln 595 600 605
         Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys Ala 610 620
         Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu Ala
625 630 635 640
10
         Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn Lys
645 650 655
         Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg Glu 660 670
15
         Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr Thr 675 680 685
         Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala Asp
690 695 700
         His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu Arg
705 710 715 720
         Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu
725 730 735
25
         Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val Gly 740 745 750
         Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg Phe 755 760 765
30
         Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser Ser 770 780
         Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp Leu Glu
785 790 800
35
         His His His His His 805
         <210>
                   144
         <211>
                   2256
40
         <212>
                   DNA
                   Artificial Sequence
          <213>
         <220>
         <223>
                   ORF46.1-961c
         <400>
                   144
45
         atgtcagatt tggcaaacga ttcttttatc cggcaggttc tcgaccgtca gcatttcgaa
                                                                                             60
                                                                                             120
         cccgacggga aataccacct attcggcagc agggggggaac ttgccgagcg cagcggccat
         atcggattgg gaaaaataca aagccatcag ttggggcaacc tgatgattca acaggcggcc
attaaaggaa atatcggcta cattgtccgc ttttccgatc acgggcacga agtccattcc
                                                                                             180
                                                                                             240
         cccttcgaca accatgcctc acattccgat tctgatgaag ccggtagtcc cgttgacgga tttagccttt accgcatcca ttgggacgga tacgaacacc atcccgccga cggctatgac
                                                                                             300
                                                                                             360
50
         gggccacagg gcggcggcta tcccgctccc aaaggcgcga gggatatata cagctacgac
                                                                                             420
         ataaaaggcg ttgcccaaaa tatccgcctc aacctgaccg acaaccgcag caccggacaa
cggcttgccg accgtttcca caatgccggt agtatgctga cgcaaggagt aggcgacgga
                                                                                             480
                                                                                             540
                                                                                             600
         ttcaaacgcg ccacccgata cagccccgag ctggacagat cgggcaatgc cgccgaagcc
         ttcaacggca ctgcagatat cgttaaaaac atcatcggcg cggcaggaga aattgtcggc
                                                                                             660
                                                                                             720
         gcaggcgatg ccgtgcaggg cataagcgaa ggctcaaaca ttgctgtcat gcacggcttg
55
         ggtctgcttt ccaccgaaaa caagatggcg cgcatcaacg atttggcaga tatggcgcaa
                                                                                             780
         čtcaaagact atgccgcagc agccatccgc gattgggcag tccaaaaccc caatgccgca
                                                                                             840
```

```
caaggcatag aagccgtcag caatatcttt atggcagcca tccccatcaa agggattgga
gctgttcggg gaaaatacgg cttgggcggc atcacggcac atcctatcaa gcggtcgcag
                                                                                                       900
                                                                                                       960
           atgögcgčőa tegeattgée gaaagggaaa teegeégtea gegacaattt tgéégatgeg
                                                                                                       1020
          gcatacgcca aatacccgtc cccttaccat tcccgaaata tccgttcaaa cttggagcag cgttacggca aagaaaacat cacctcctca accgtgccgc cgtcaaacgg caaaaatgtc
                                                                                                       1080
                                                                                                       1140
          aāactggcag accaacgcca cccgaagaca ggcgtaccgt ttgacggtaa agggtttccg
                                                                                                       1200
          aattttgaga agcacgtgaa atatgatacg ggatccggag gaggaggagc cacaaacgac
gacgatgtta aaaaagctgc cactgtggcc attgctgctg cctacaacaa tggccaagaa
                                                                                                       1260
1320
          ātcāacīgtt tcaaagctīg agagāccātc tacīgacāttīg atgaagacīg cacaattācc
                                                                                                       1380
           aaaaaagacg caactgcagc cgatgttgaa gccgacgact ttaaaggtct gggtctgaaa
                                                                                                       1440
           aaagtogtgā ctaacotgāc cāaaācogto āatgaaāaca aacaaāācgt cgātgocaaa
                                                                                                       1500
10
           gtaãaagctg cagaatctga aatagaaãag ttaãcaacca agttagcaga cáctgatgcc
                                                                                                       1560
           gctttagcag atactgatgc cgctctggat gcaaccacca acgccttgaa taaattggga
                                                                                                       1620
          gaaaatataa cgacatttgc tgaagagact aagacaaata tcgtaaaaat tgatgaaaaa ttagaagccg tggctgatac cgtcgacaag catgccgaag cattcaacga tatcgccgat tcattggatg aaaccaacac taaggcagac gaagccgtca aaaccgccaa tgaagccaaa
                                                                                                       1680
                                                                                                       1740
                                                                                                       1800
          cagacggccg aagaaaccaa acaaaacgtc gatgccaaag taaaagctgc agaaactgca
gcaggcaaag ccgaagctgc cgctggcaca gctaatactg cagccgacaa ggccgaagct
                                                                                                       1860
                                                                                                       1920
           gtcgctgcaā aagttāccga cātcāāagct gatatcgctā cgāacāaaga tāatāttgct
                                                                                                       1980
          aaaaaagcaa acagtgccga cgtgtacacc agagaagagt ctgacagcaa atttgtcaga attgatggtc tgaacgctac taccgaaaaa ttggacacac gcttggcttc tgctgaaaaa
                                                                                                       2040
                                                                                                       2100
           tccattocco atcacoatac tcocctoaac gotttogata aaacagtotc agacctococ
                                                                                                       2160
          aaagaaaccc gccaaggcct tgcagaacaa gccgcgctct ccggtctgtt ccaaccttac aacgtgggtc tcgagcacca ccaccaccac cactga
                                                                                                       2220
2256
20
           <210>
                     145
                     751
          <211>
           <212>
           <213>
                     Artificial Sequence
25
          <220>
           <223>
                     ORF46.1-961c
          Met Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg
1 5 10 15
30
          Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg Gly 20 25 30
          Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln Ser
35 40
35
          His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly Asn 50 55 60
          Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu Val His Ser 65 70 75 80
40
          Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly Ser 85 90 95
          Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu 100 105
          His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Gly Tyr Pro
115 120 125
          Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val
130 140
50
          Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln 145 150 155
          Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly 175
          Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp 180 \hspace{1cm} 185 \hspace{1cm} 190
```

Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val 195 200 205 Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp Ala 210 220 Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu 225 230 235 240 Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala 245 250 255 10 Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ala Ile Arg Asp Trp 260 265 270 Ala Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn 275 280 285 15 Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly
290 295 300 Lys Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln 305 310 315 20 Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn 325 330 335 Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg 340 345 350 25 Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr 355 360 365 Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp 370 380 30 Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro 385 390 395 400 Asn Phe Glu Lys His Val Lys Tyr Asp Thr Gly Ser Gly Gly Gly Gly 415 35 Ala Thr Asn Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile Ala 420 425 430 Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly Glu 435 440 445 40 Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp Ala 450 455 460 Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu Lys 465 470 480 45 Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln Asn 485 490 495 Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu Thr 500 510 50 Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala Ala 515 520 525 Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile Thr 530 540 Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu Lys 55

```
Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe Asn 565 570
            Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu Ala
580 585 590
            Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys Gln
595 600 605
            Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys Ala
610 615 620
10
            Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu Ala
625 630 635
            Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn Lys
645 650 655
15
            Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg Glu 660 670
            Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr Thr
675 680 685
20
            Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala Asp
690 695 700
            His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu Arg
705 710 715 720
25
            Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu
725 730 735
            Phe Gln Pro Tyr Asn Val Gly Leu Glu His His His His His 740 745 750
30
            <210>
                        146
            <211>
                        2421
            <212>
                        DNA
                        Artificial Sequence
            <213>
35
            <220>
                        961-ORF46.1
            <223>
            <400>
                        146
            atggccacaa acgacgacga tgttaaaaaaa gctgccactg tggccattgc tgctgcctac
                                                                                                                    60
                                                                                                                    120
            aacaatggcc aagaaatcaa cggtttcaaa gctggagaga ccatctacga cattgatgaa
            gacggcacaa ttaccaaaaa agacgcaact gcagccgatg ttgaagccga cgactttaaa ggtctgggtc tgaaaaaagt cgtgactaac ctgaccaaaa ccgtcaatga aaacaaacaa aacgtcgatg ccaaagtaaa agctgcagaa tctgaaatag aaaagttaac aaccaagtta gcagacactg atgccgcttt agcagatact gatgccgctc tggatgcaac caccaacgcc
40
                                                                                                                    180
                                                                                                                     240
                                                                                                                     300
                                                                                                                     360
            ttgaataaat tgggagaaaa tataacgaca tttgctgaag agactaagac aaatatcgta aaaattgatg aaaaattaga agccgtggct gataccgtcg acaagcatgc cgaagcattc
                                                                                                                     420
                                                                                                                    480
            aacgatatcg ccgattcatt ggatgaaacc aacactaagg cagacgaagc cgtcaaaacc
                                                                                                                    540
45
            gccaatgaag ccaaacagac ggccgaagaa accaaacaaa acgtcgatgc caaagtaaaa
gctgcagaaa ctgcagcagg caaagccgaa gctgccgctg gcacagctaa tactgcagcc
gacaaggccg aagctgtcgc tgcaaaagtt accgacatca aagctgatat cgctacgaac
                                                                                                                    600
                                                                                                                    660
                                                                                                                     720
            aaagataata ttgctaaaaa agcaaacagt gccgacgtgt acaccagaga agagtctgac agcaaatttg tcagaattga tggtctgaac gctactaccg aaaaaattgga cacacgcttg gcttctgctg aaaaatccat tgccgatcac gatactcgcc tgaacggttt ggataaaaca
                                                                                                                    780
                                                                                                                    840
                                                                                                                    900
50
            gtgtcagacc tgcgcaaaga aacccgccaa ggccttgcag aacaagccgc gctctccggt
ctgttccaac cttacaacgt gggtcggttc aatgtaacgg ctgcagtcgg cggctacaaa
tccgaatcgg cagtcgccat cggtaccggc ttccgcttta ccgaaaactt tgccgccaaa
                                                                                                                    960
                                                                                                                    1020
                                                                                                                    1080
            gcaggcgtgg cagtcggcac ttcgtccggt tcttccgcag cctaccatgt cggcgtcaat
                                                                                                                    1140
            tacgagtggg gatccggagg aggaggatca gatttggcaa acgattctt tatccggcaggttctcgacc gtcagcattt cgaacccgac gggaaatacc acctattcgg cagcaggggg
                                                                                                                    1200
                                                                                                                    1260
55
            gaacttgccg agcgcagcgg ccatatcgga ttgggaaaaa tacaaagcca tcagttgggc
                                                                                                                    1320
            aacctgatga ticaacaggc ggccattaaa ggaaatatcg gctacattgt ccgcttttcc
                                                                                                                    1380
```

```
gatcacgggc acgaagtcca ttcccccttc gacaaccatg cctcacattc cgattctgat
                                                                                                                1440
                                                                                                                1500
          gaagccggta gtčccgttga cggatttagc čtttaccgca tccattggga cggatacgaa
          caccatcccg ccgacggcta tgacgggcca cagggcggcg gctatcccgc tcccaaaggc
gcgagggata tatacagcta cgacataaaa ggcgttgccc aaaatatccg cctcaacctg
                                                                                                                1560
                                                                                                                1620
           accgacaacc gcagcaccgg acaacggctt gccgaccgtt tccacaatgc cggtagtatg
                                                                                                                1680
          ctgacgcaag gagtaggcga cggattcaaa cgcgccaccc gatacagccc cgagctggac agatcggca atgccgccga agccttcaac ggcactgcag atatcgttaa aaacatcatc
                                                                                                                1740
                                                                                                                1800
          agattgggta atgccgcga agctttaat ggcattgtay atattgtta adacattatt
ggcgcggcag gagaaattgt cggcgcaggc gatgccgtgc agggcataag cgaaggctca
aacattgctg tcatgcacgg cttgggtctg ctttccaccg aaaacaagat ggcgcgcatc
aacgatttgg cagatatggc gcaactcaaa gactatgccg cagcagccat ccgcgattgg
gcagtccaaa accccaatgc cgcacaaggc atagaagccg tcagcaatat ctttatggca
gccatcccca tcaaagggat tggagctgtt cgggaaaat acgcttggg cggcatcacc
                                                                                                                1860
                                                                                                                1920
                                                                                                                1980
                                                                                                                2040
10
                                                                                                                2100
          gcacatecta teaageggte geagatggge gegategeat tgeegaaagg gaaateegee
gteagegaea attttgeega tgeggeatae geeaaataee egteeeetta eeatteeega
                                                                                                                2160
                                                                                                                2220
                                                                                                                2280
           aatatccgtt caaacttgga gcagcgttac ggcaaagaaa acatcacctc ctcaaccgtg
          ccgccgtcaa acggcaaaaa tgtcaaactg gcagaccaac gccacccgaa gacaggcgta ccgtttgacg gtaaagggtt tccgaatttt gagaagcacg tgaaatatga tacgctcgag
                                                                                                                2340
                                                                                                                2400
           caccaccacc accaccactg a
                                                                                                                2421
           <210>
                      147
           <211>
                      806
           <212>
                      PRT
                      Artificial Sequence
           <213>
20
           <220>
           <223>
                      961-ORF46.1
          Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile
1 5 10 15
25
          Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly 20 25 30
          Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp 35 40 45
30
          Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu 50 60
           Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln 65 70 75 80
35
          Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu
85 90 95
          Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala
100 105 110
40
          Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile
115 120 125
          Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu 130 135 140
45
                Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe
150 155 160
          Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu
165 170 175
50
          Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys 180 \hspace{1cm} 185 \hspace{1cm} 190
          Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys
195 200 205
          Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu
210 215 220
```

Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn 225 230 240 Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg 245 250 255 Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr 260 265 270 Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala 275 280 285 10 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu 290 295 300 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly 305 310 315 320 15 Leu Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val 325 330 335 Gly Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg  $340 \hspace{1.5cm} 345 \hspace{1.5cm} 350$ 20 Phe Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser 355 360 365 Ser Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp Gly 370 375 380 25 Ser Gly Gly Gly Gly Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln 385 390 400 Val Leu Asp Arg Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe 405 410 41530 Gly Ser Arg Gly Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly 420 425 Lys Ile Gln Ser His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala 445 35 Ile Lys Gly Asn Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His 450 460 Glu Val His Ser Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp 465 470 475 480 40 Glu Ala Gly Ser Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp 485 490 495 Asp Gly Tyr Glu His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly 500 50545 Gly Gly Tyr Pro Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp 515 525 Ile Lys Gly Val Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg 530 540 50 Ser Thr Gly Gln Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met 545 550 560 Leu Thr Gln Gly Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser 570 575 Pro Glu Leu Asp Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr 580 585 590

```
Ala Asp Ile Val Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly 595 600 605
          Ala Gly Asp Ala Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val
610 620
           Met His Gly Leu Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile
625 630 635
          Asn Asp Leu Ala Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ala 655 655
10
           Ile Arg Asp Trp Ala Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu
660 665 670
          Ala Val Ser Asn Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly 675 680 685
15
          Ala Val Arg Gly Lys Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile
690 700
           Lys Arg Ser Gln Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala
705 710 715 720
20
          Val Ser Asp Asn Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro 725 730 735
          Tyr His Ser Arg Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys 740 745 750
25
          Glu Asn Ile Thr Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val 755 760 765
          Lys Leu Ala Asp Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly 770 780
30
           Lys Gly Phe Pro Asn Phe Glu Lys His Val Lys Tyr Asp Thr Leu Glu
785 790 795 800
           His His His His His
35
           <210>
                     148
           <211>
                     1938
           <212>
                     DNA
           <213>
                     Artificial Sequence
           <220>
40
           <223>
                     961-741
           <400>
                     148
           atggccacaa acgacgacga tgttaaaaaa gctgccactg tggccattgc tgctgcctac
                                                                                                      60
          aacaatggcc aagaaatcaa cggtttcaaa gctggagaga ccatctacga cattgatgaa
                                                                                                      120
                                                                                                      180
           gacggcacaa ttaccaaaaa agacgcaact gcagccgatg ttgaagccga cgactttaaa
45
          ggtctgggtc tgaaaaaagt cgtgactaac ctgaccaaaa ccgtcaatga aaacaaacaa aacgtcgatg ccaaagtaaa agctgcagaa tctgaaatag aaaagttaac aaccaagtta gcagacactg atgccgctt agcagatact gatgccgctc tggatgcaac caccaacgcc
                                                                                                      240
                                                                                                      300
                                                                                                      360
          ttgaataaat tgggagaaaa tataacgaca tttgctgaag agactaagac aaatatcgta
aaaattgatg aaaaattaga agccgtggct gataccgtcg acaagcatgc cgaagcattc
                                                                                                      420
                                                                                                      480
          aacgatatcg ccgattcatt ggatgaaacc aacactaagg cagacgaagc cgtcaaaacc
                                                                                                      540
50
          gccaatgaag ccaaacagac ggccgaagaa accaaacaaa acgtcgatgc caaagtaaaa gctgcagaaa ctgcagcagg caaagccgaa gctgccgctg gcacagctaa tactgcagcc gacaaggccg aagctgtcgc tgcaaaagtt accgacatca aagctgatat cgctacgaac
                                                                                                      600
                                                                                                      660
                                                                                                      720
          aaagataata ttgctaaaaa agcaaacagt gccgacgtgt acaccagaga agagtctgac
agcaaatttg tcagaattga tggtctgaac gctactaccg aaaaattgga cacacgcttg
                                                                                                      780
                                                                                                      840
          gcttctgctg aaaaatccat tgccgatcac gatactcgcc tgaacggttt ggataaaaca
                                                                                                      900
55
          gtgtcagacc tgcgcaaaga aacccgccaa ggccttgcag aacaagccgc gctctccggt
                                                                                                      960
          ctgttccaac citacaacgt gggtcggttc aatgtaacgg ctgcagtcgg cggctacaaa
                                                                                                      1020
```

1080

```
tccgaatcgg cagtcgccat cggtaccggc ttccgcttta ccgaaaactt tgccgccaaa
           gcağgcgtőg cagtcógcac ticgtccógt tettécgcag ceiaccatgt cógcótcaat
                                                                                                   1140
           tacgagtggg gatccggagg gggtggtgtc gccgccgaca tcggtgcggg gcttgccgat
gcactaaccg caccgctcga ccataaagac aaaggtttgc agtctttgac gctggatcag
                                                                                                   1200
                                                                                                   1260
           tccgtcagga aaaacgagaa actgaagctg gcggcacaag gtgcggaaaa aacttatgga aacggtgaca gcctcaatac gggcaaattg aagaacgaca aggtcagccg tttcgacttt atccgccaaa tcgaagtgga cgggcagctc attaccttgg agagtggaga gttccaagta
                                                                                                   1320
                                                                                                   1380
                                                                                                   1440
                                                                                                   1500
           tacaāacaaa gccatīccgc cīīāaccgcc tttcagaccg agcāaāīaca āgattcggag
           cattccggga agatggttgc gaaacgccag ttcagaatcg gcgacatagc gggcgaacat acatcttttg acaagcttcc cgaaggcggc agggcgacat atcgcgggac ggcgttcggt
                                                                                                   1560
                                                                                                   1620
10
           tcagacgatg ccggcggaaa actgacctac accatagatt tcgccgccaa gcagggaaac
                                                                                                   1680
           ggcăaaătcg aacăttigaa atcgccagaa ctcaatgtcg acctggccgc cgccgatatc
aagccggatg gaaaacgcca tgccgtcatc agcggttccg tcctttacaa ccaagccgag
                                                                                                   1740
                                                                                                   1800
           aaaggčagtť actcccťcgg tatcťttggc ggaaaagccc aggaagttgc cggcagcgcg
                                                                                                   1860
           gaagīgaāaa cegtaaacģģ cataegecāt āīeggecīttg ceģecāagēa acīegāgēae
                                                                                                   1920
                                                                                                   1938
           caccaccacc accactga
15
           <210>
                     149
                     645
           <211>
           <212>
                     PRT
           <213>
                     Artificial Sequence
           <220>
20
                     961-741
           <223>
           <400>
           Met Ala Thr Asn Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile
1 5 10 15
25
           Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly 20 25 30
           Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp 40 45
           Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu 50 60
30
           Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln 65 70 75 80
           Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu
85 90 95
35
           Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala 100 105
           Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile
115 120
40
           Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu 130 140
           Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe
145 150 155 160
45
           Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu 165 170 175
           Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys 180 185 190
50
           Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys
195 200 205
           Ala Glu Ala Ala Gly Thr Ala Asn Thr Ala Asp Lys Ala Glu
210 215 220
55
           Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn
```

	225					230					235					240
5	Lys	Asp	Asn	Ile	Ala 245	Lys	Lys	Ala	Asn	ser 250	Ala	Asp	٧a٦	Tyr	Thr 255	Arg
	Glu	Glu	Ser	Asp 260	Ser	Lys	Phe	Val	Arg 265	Ile	Asp	Gly	Leu	Asn 270	Ala	Thr
10	Thr	Glu	Lys 275	Leu	Asp	Thr	Arg	Leu 280	Ala	Ser	Ala	Glu	Lys 285	Ser	Ile	Ala
	Asp	Ніs 290	Asp	Thr	Arg	Leu	Asn 295	Gly	Leu	Asp	Lys	Thr 300	val	Ser	Asp	Leu
15	Arg 305	Lys	Glu	Thr	Arg	G]n 310	Gly	Leu	Ala	Glu	Gln 315	Аlа	Ala	Leu	Ser	Gly 320
	Leu	Phe	Gln	Pro	Tyr 325	Asn	∨al	Gly	Arg	Phe 330	Asn	۷a٦	Thr	Ala	Ala 335	۷al
00	Gly	Gly	Tyr	Lys 340	Ser	Glu	Ser	Ala	Va1 345	Ala	Ile	Gly	Thr	G]y 350	Phe	Arg
20	Phe	Thr	G]u 355	Asn	Phe	Аlа	Ala	Lys 360	Аlа	Gly	val	Ala	va1 365	Gly	Thr	ser
	Ser	Gly 370	Ser	Ser	Ala	Ala	Tyr 375	His	<b>∨</b> a1	Gly	∨al	Asn 380	туг	Glu	Тгр	Gly
25	Ser 385	Gly	Gly	Gly	Gly	Va1 390	Ala	Ala	Asp	Ile	Gly 395	Ala	Gly	Leu	Ala	Asp 400
	Ala	Leu	Thr	Ala	Pro 405	Leu	Asp	ніѕ	Lys	Asp 410	Lys	Gly	Leu	Gln	Ser 415	Leu
30	Thr	Leu	Asp	G1n 420	Ser	val	Arg	Lys	Asn 425	Glu	Lys	Leu	Lys	Leu 430	Ala	Ala
	Gln	Gly	Ala 435	Glu	Lys	Thr	Tyr	Gly 440	Asn	Gly	Asp	Ser	Leu 445	Asn	Thr	Gly
35	Lys	Leu 450	Lys	Asn	Asp	Lys	Val 455	Ser	Arg	Phe	Asp	Phe 460	Ile	Arg	Gln	Ile
	Glu 465	Val	Asp	Gly	Gln	Leu 470	Ile	Thr	Leu	Glu	Ser 475	Gly	Glu	Phe	Gln	Val 480
40					485	Ser				490					495	
	Gln	Asp	ser	G1u 500	His	Ser	Gly	Lys	Met 505	٧a٦	Ala	Lys	Arg	Gln 510	Phe	Arg
45	Ile	Gly	Asp 515	Ile	Ala	Gly	Glu	ніs 520	Thr	Ser	Phe	Asp	Lys 525	Leu	Pro	Glu
	Gly	G]y 530	Arg	Ala	Thr	Tyr	Arg 535	Gly	Thr	Ala	Phe	Gly 540	Ser	Asp	Asp	Ala
50	Gly 545	Gly	Lys	Leu	Thr	Tyr 550	Thr	Ile	Asp	Phe	Ala 555	Ala	Lys	Gln	Gly	Asn 560
	Gly	Lys	Ile	Glu	нis 565	Leu	Lys	Ser	Pro	G]u 570	Leu	Asn	val	Asp	Leu 575	Ala
55	Ala	Ala	Asp	11e 580	Lys	Pro	Asp	Gly	Lys 585	Arg	ніѕ	Ala	٧a٦	11e 590	Ser	Gly
	ser	٧a٦	Leu	Tyr	Asn	Gln	Ala	Glu	Lys	Gly	Ser	Tyr	Ser	Leu	Gly	Ile

```
595
                                                                                                             600
                                                                                                                                                                        605
                          Phe Gly Gly Lys Ala Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr
610 615 620
                          Val Asn Gly Ile Arg His Ile Gly Leu Ala Ala Lys Gln Leu Glu His
                           His His His His His
10
                           <210>
                                                  150
                                                  4335
                           <211>
                           <212>
                                                  DNA
                                                 Artificial Sequence
                           <213>
                           <220>
15
                                                 961-983
                           <223>
                           <400>
                                                  150
                           atggccacaa acgacgacga tgttaaaaaa gctgccactg tggccattgc tgctgcctac
                                                                                                                                                                                                                                   120
                           aacaatggcc aagaaatcaa cggtttcaaa gctggagaga ccatctacga cattgatgaa
                          gacggcacaa ttaccaaaaa agacgcaact gcagccgatg ttgaagccga cgactttaaa ggtctgggtc tgaaaaaagt cgtgactaac ctgaccaaaa ccgtcaatga aaacaaacaa
                                                                                                                                                                                                                                   180
20
                                                                                                                                                                                                                                   240
                          aacgtcgatg ccaaagtaaa agctgcagaa tctgaaatag aaaagttaac aaccaagtta gcagacactg atgccgctt agcagatact gatgccgctc tggatgcaac caccaacgcc ttgaataaat tgggagaaaa tataacgaca tttgctgaag agactaagac aaatatcgta
                                                                                                                                                                                                                                   300
                                                                                                                                                                                                                                   360
                                                                                                                                                                                                                                   420
                          aaaattgatg aaaaattaga agccgtggct gataccgtcg acaagcatgc cgaagcattc aacgatatcg ccgattcatt ggatgaaacc aacactaagg cagacgaagc cgtcaaaacc gccaatgaag ccaaacagac ggccgaagaa accaaacaaa acgtcgatgc caaagtaaaa
                                                                                                                                                                                                                                   480
                                                                                                                                                                                                                                   540
25
                                                                                                                                                                                                                                   600
                         gccaatgaag ccaaacagac ggccgaagaa accaaacaaa acgtcgatgc caaagtaaaa gctgcagaaaa ctgcagcagg caaagccgaa gctgccgctg gcacaagctaa tactgcagcc gacaaggccg aagctgtcgc tgcaaaagtt accgacatac aagctgatat cgctacgaac aaagataata ttgctaaaaa agcaaacagt gccgacgtgt acaccaggag agagtctgac agcaaatttg tcagaattga tggtctgaac gctactaccg aaaaattgga cacacgcttg gcttctgctg aaaaatccat tgccgatcac gatactcgcc tgaacggttt ggataaaaca gtgtcagacc tgcgcaaaga aacccgccaa ggccttgcag aacaaagccgc gctctccggt ctgttccaac cttacaacgt gggtcggttc aatgtaacgg ctgcagtcgg cggctacaaa tccgaatcgg cagtcggcac ttcgccggt tctccccag cctaccatgt cggcgcaaa gaggcggtgg cagtcggcac ttcgccggt tctcccca acctacatgt cggcgcaaa gaggcggtgg gatccggcac ttctgcccga acttcaatgc aggcggtacca
                                                                                                                                                                                                                                   660
                                                                                                                                                                                                                                   720
                                                                                                                                                                                                                                   780
                                                                                                                                                                                                                                   840
                                                                                                                                                                                                                                   900
                                                                                                                                                                                                                                   960
30
                                                                                                                                                                                                                                   1020
                                                                                                                                                                                                                                   1080
                                                                                                                                                                                                                                   1140
                          tacgagtggg gatccggcgg aggcggcact tctgcgcccg acttcaatgc aggcggtacc
ggtatcggca gcaacagcag agcaacaaca gcgaaatcag cagcagtatc ttacgccggt
atcaagaacg aaatgtgcaa agacagaagc atgctctgtg ccggtcggga tgacgttgcg
                                                                                                                                                                                                                                   1200
                                                                                                                                                                                                                                   1260
                                                                                                                                                                                                                                   1320
35
                           gttacagaca gggatgccaa aatcaatgcc cccccccga atctgcatac cggagacttt
                                                                                                                                                                                                                                   1380
                          ccaaacccaa atgacgcata caagaatttg atcaacctca aacctgcaat tgaagcaggc
tatacaggac gcggggtaga ggtaggtatc gtcgacacag gcgaatccgt cggcagcata
                                                                                                                                                                                                                                   1440
                                                                                                                                                                                                                                   1500
                           tcctttcccg aactgtatgg cagaaaagaa cacggctata acgaaaatta caaaaactat
                                                                                                                                                                                                                                   1560
                          acggcgtata tgcggaagga agcgcctgaa gacggaggcg gtaaagacat tgaagcttct ttcgacgatg aggccgttat agagactgaa gcaaagccga cggatatccg ccacgtaaaa gaaatcggac acatcgattt ggtctcccat attattggcg ggcgttccgt ggacggcaga cctgcaggcg gtattgcgc cgatgcgacg ctacacataa tgaatacgaa tgatggaacc aagaacgaaa tgatggttgc agccatcgc aatgcatggg tcaagctggg cgaacgtggc
                                                                                                                                                                                                                                   1620
                                                                                                                                                                                                                                   1680
40
                                                                                                                                                                                                                                   1740
                                                                                                                                                                                                                                   1800
                                                                                                                                                                                                                                   1860
                          gtgcgcatcg tcaataacag ttttggaaca acatcgaggg caggcactgc cgaccttttc caaatagcca attcggagga gcagtaccgc caagcggtgc tcgactattc cggcggtgat aaaacagacg agggtatccg cctgatgcaa cagagcgatt acggcaacct gtcctaccac atccgtaata aaaacatgct tttcatcttt tcgacaggca atgacgcaca agctcagccc
                                                                                                                                                                                                                                   1920
                                                                                                                                                                                                                                   1980
                                                                                                                                                                                                                                   2040
                                                                                                                                                                                                                                   2100
45
                          accacataty ccctattycc attitatett tegacayyca atyacycaca ageteayett accacataty ccctattycc attitateaa aaagacgete aaaaaggeat tateacayte geaggegtag accgeaytyg agaaaagtte aaacgggaaa tytatygaga accygytaca gaaccgetty agtatygete caaccattyc ggaattacty ccatytygty cctyteggea ccctatygaay caagegteey titeacceyt acaaacceya tieaaattyc eggaacatee titteegeac ccategtaac eggeacggey getetyetye tyeagaaata eccytygaty ageaacgaca acctycytac caegttyety acgaegyete aggacategy tyeagtegge
                                                                                                                                                                                                                                   2160
                                                                                                                                                                                                                                   2220
                                                                                                                                                                                                                                   2280
                                                                                                                                                                                                                                   2340
                                                                                                                                                                                                                                   2400
                                                                                                                                                                                                                                    2460
                         agcaacgaca acctgcgtac cacgttgctg acgacggctc aggacatcgg tgcagtcggc gtggacagca agttcggctg gggactgctg gatgcgggta aggccatgaa cggacccgcg tcctttccgt tcggcgactt taccgccgat acgaaaggta catccgatat tgcctactcc ttccgtaacg acattcagg cacggcggc ctgatcaaaa aaggcggcag ccaactgcaa ctgcacggca acaacaccta tacgggcaaa accattatcg aaggcggttc gctggtgttg tacggcaaca acaaatcgga tatgcgcgtc gaaaccaaag gtgcgctgat ttataacggg gcggcatccg gcggcagcct gaacagcgac ggcattgtct atctggcaga taccgaccaa tccggcgcaa acgaaaccgt acacatcaaa ggcagtctgc agctggacgg caaaggtacg ctgtacacac gtttgggcaa actgctgaaa gtggacggta cggcggcatac cggcggcaag
50
                                                                                                                                                                                                                                   2520
                                                                                                                                                                                                                                    2580
                                                                                                                                                                                                                                    2640
                                                                                                                                                                                                                                   2700
                                                                                                                                                                                                                                   2760
                                                                                                                                                                                                                                   2820
55
                                                                                                                                                                                                                                   2880
                                                                                                                                                                                                                                   2940
```

3600 3660

3720

4020 4080

4140 4200 4260

4320

4335

```
ctgtacatgt cggcacgcgg Caagggggca ggctatctca acagtaccgg acgacgtgtt cccttcctga gtgccgccaa aatcgggcag gattattctt tcttcacaaa catcgaaacc gacggcggcc tgctggcttc cctcgacagc gtcgaaaaaa cagcgggcag tgaaggcgac acgctgtcct attatgtccg tcgcggcaat gcggcacgga ctgcttcggc agcggcacat tccgcgcccg ccggtctgaa acacgccgta gaacagggcg gcagcaatct ggaaaacctg atggtcgaac tggatgcctc cgaatcatcc gcaacacccg agacggttga aactgcggca gccgaccgca cagatatgcc gggcatccgc ccctacggcg caactttccg cgcagcggca gccgtacagc atgcgaatgc cgccgacggt gtacgcatct tcaacagtct cgccgctacc gtctatgccg acagtaccgc cgccaagcc gatatgcagg gacgcgcct gaaagccgta tcggacgggt tggaccacaa cggcacgggt ctgcgcgtca tcgcgcaaac ccaacaggac ggtggaacgt gggaacagt ccaaaccgtc
                  ggtggaacgt gggaacaggg cggtgttgaa ggcaaaatgc gcggcagtac ccaaaccgtc
ggcattgccg cgaaaaccgg cgaaaatacg acagcagccg ccacactggg catgggacgc
10
                  agcacatgga gcgaaaacag tgcaaatgca aaaaccgaca gcattagtct gtttgcaggc
                  atacggcacg atgcgggcga tatcggctat ctcaaaggcc tgttctccta cggacgctac aaaaacagca tcagccgcag caccggtgcg gacgaacatg cggaaggcag cgtcaacggc acgctgatgc agctgggcgc actgggcggt gtcaacgttc cgtttgccgc aacgggagat ttgacggtcg aaggcggtct gcgctacgac ctgctcaaac aggacgctgt cgccgaaaaa ggcagtgctt tgggctggag cggcaacagc ctcactgaag gcacgctggt cggactcgcg ggtctgaagc tgccgcaacc ctgagcgat aaagccgtcc tgtttgcaac ggcggggtg
15
                  gaacgcgacc tgaacggacg cgactacacg gtaacgggcg gctttaccgg cgcgactgca
gcaaccggca agacggggc acgcaatatg ccgcacaccc gtctggttgc cggcctgggc
gcggatgtcg aattcggcaa cggctggaac ggcttggcac gttacagcta cgccggttcc
                  aaačagtacg gcaaccacag cggacgagtc ggcgtaggct accggttcct cgagcaccac caccaccacc actga
20
                  <210>
                                    151
                                    1444
                  <211>
                   <212>
                                   PRT
                                   Artificial Sequence
                  <220>
                                   961-983
                  <223>
                  Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile
1 5 10 15
30
                  Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly 20 25 30
                  Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp 45
35
                  Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu 50 60
                  Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln 65 70 75 80
40
                  Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu
85 90 95
                  Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala
100 105 110
45
                  Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile
115 120 125
                  Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu 130 140
50
                  Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe
145 150 155 160
                  Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu
165 170 175
                  Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys
180 185 190
```

Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys 195 200 205 Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu 210 215 220 Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn 225 230 240 Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg 245 250 255 10 Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr 260 265 270 Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala 275 280 285 15 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu 290 295 300 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly 305 310 315 320 20 Leu Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val 325 330 335 Gly Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg 340 345 35025 Phe Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser 355 360 365 Ser Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp Gly 370 380 30 Ser Gly Gly Gly Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr 385 390 395 400 Gly Ile Gly Ser Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val 405 410 415 35 Ser Tyr Ala Gly Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu 420 425 430 Cys Ala Gly Arg Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile 435 440 445 40 Asn Ala Pro Pro Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn 450 460 Asp Ala Tyr Lys Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly 465 470 480 45 Tyr Thr Gly Arg Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser 485 490 495 Val Gly Ser Ile Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly 500 510 50 Tyr Asn Glu Asn Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala 515 520 525 Pro Glu Asp Gly Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu 530 540 Ala Val Ile Glu Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys 545 550 560 55

Glu Ile Gly His Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser 570 575 Val Asp Gly Arg Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His 580 585 590 Ile Met Asn Thr Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala 595 600 605 Ile Arg Asn Ala Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val 610 615 620 10 Asn Asn Ser Phe Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe 625 630 635 Gln Ile Ala Asn Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr 645 650 655 15 Ser Gly Gly Asp Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser 660 670 Asp Tyr Gly Asn Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe 675 680 685 20 Ile Phe Ser Thr Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala 690 695 700 Leu Leu Pro Phe Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val 705 710 715 720 25 Ala Gly Val Asp Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly
725 730 735 Glu Pro Gly Thr Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile 740 745 750 30 Thr Ala Met Trp Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe 755 760 765 Thr Arg Thr Asn Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro 770 775 780 35 Ile Val Thr Gly Thr Ala Ala Leu Leu Gln Lys Tyr Pro Trp Met 785 790 800 Ser Asn Asp Asn Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile 805 810 81540 Gly Ala Val Gly Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala 820 825 830 Gly Lys Ala Met Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr 835 840 845 45 Ala Asp Thr Lys Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp 850 860 Ile Ser Gly Thr Gly Gly Leu Ile Lys Lys Gly Gly Ser Gln Leu Gln 865 870 87550 Leu His Gly Asn Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly 885 890 895 Ser Leu Val Leu Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr 900 905 910 55 Lys Gly Ala Leu Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn 915 925

	Ser	930	GIY	IIe	vaı	ıyr	935	АІа	Asp	Thr	Asp	940	Ser	GIY	АІа	ASN
5	G]u 945	Thr	٧a٦	His	Ile	Lys 950	Gly	Ser	Leu	Gln	Leu 955	Asp	Glу	Lys	Gly	Thr 960
	Leu	Tyr	Thr	Arg	Leu 965	Glу	Lys	Leu	Leu	Lys 970	٧a٦	Asp	Gly	Thr	Ala 975	Ile
10	Ile	Gly	Gly	Lys 980	Leu	Tyr	Met	Ser	Ala 985	Arg	Gly	Lys	Gly	Ala 990	Gly	Tyr
	Leu	Asn	Ser 995	Thr	Gly	Arg	Arg	Val 1000	Pro )	Phe	Leu	Ser	Ala 1005		Lys	Ile
15	Gly	Gln 1010		Tyr	Ser	Phe	Phe 1015		Asn	Ile	Glu	Thr 1020		Gly	Gly	Leu
	Leu 1025		ser	Leu	Asp	Ser 1030		Glu	Lys	Thr	Ala 1035		Ser	Glu	Gly	Asp 1040
20	Thr	Leu	Ser	Tyr	Tyr 1045		Arg	Arg	Gly	Asn 1050		Ala	Arg	Thr	Ala 1055	
	Ala	Ala	Ala	ніs 1060		Ala	Pro	Ala	Gly 1065		Lys	His	Ala	Val 1070		Gln
25	GТу	Gly	Ser 1075		Leu	Glu	Asn	Leu 1080	Met )	val	Glu	Leu	Asp 1085		Ser	Glu
	Ser	ser 1090		Thr	Pro	Glu	Thr 1095		Glu	Thr	Ala	Ala 1100		Asp	Arg	Thr
30	Asp 1105		Pro	Gly	Ile	Arg 111(		⊤yr	Gly	Ala	Thr 1115		Arg	Ala	Ala	Ala 1120
	Ala	val	G∏n	His	Ala 1125		Ala	Ala	Asp	Gly 1130		Arg	Ile	Phe	Asn 1135	
35	Leu	Ala	Ala	Thr 1140		Tyr	Ala	Asp	Ser 1145		ΑΊа	Ala	His	Ala 1150		Met
	Gln	Gly	Arg 1155		Leu	Lys	ΑΊа	val 1160	ser )	Asp	Gly	Leu	Asp 1165		Asn	Gly
40	Thr	Gly 1170		Arg	٧a٦	Ile	Ala 1175		Thr	Gln	Gln	Asp 1180		Gly	Thr	Trp
40	Glu 1185		Gly	Gly	٧a٦	Glu 1190		Lys	Met	Arg	Gly 1195		Thr	Gln	Thr	∨a1 1200
	Gly	Ile	Ala	Ala	Lys 1205		GТу	Glu	Asn	Thr 1210		ΑΊα	ΑΊа	Ala	Thr 1215	
45	Gly	Met	Gly	Arg 1220		Thr	Тгр	ser	Glu 1225		ser	ΑΊа	Asn	Аlа 1230		Thr
	Asp	Ser	Ile 1235		Leu	Phe	Ala	Gly 1240	Ile )	Arg	His	Asp	Ala 1245		Asp	Ile
50	Gly	Tyr 1250		Lys	Gly	Leu	Phe 1255		Tyr	Gly	Arg	Tyr 1260		Asn	Ser	Ile
	Ser 1265		Ser	Thr	Gly	Ala 1270		Glu	His	Аlа	Glu 1275		Ser	val	Asn	Gly 1280
55	Thr	Leu	Met	Gln	Leu 1285		Ala	Leu	Gly	Gly 1290		Asn	٧a٦	Pro	Phe 1295	

```
Ala Thr Gly Asp Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu
1300 1305 1310
               Lys Gln Asp Ala Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly 1315 1320 1325
               Asn Ser Leu Thr Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu
1330 1335 1340
                Ser Gln Pro Leu Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val
1345 1350 1355 136
10
               Glu Arg Asp Leu Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr
1365 1370 1375
               Gly Ala Thr Ala Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His
1380 1385 1390
               Thr Arg Leu Val Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly 1395 1400 1405
               Trp Asn Gly Leu Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly 1410 1420
               Asn His Ser Gly Arg Val Gly Val Gly Tyr Arg Phe Leu Glu His His
1425 1430 1435 144
                His His His His
25
                                152
2256
                <210>
                <211>
                <212>
                                DNA
                <213>
                                Artificial Sequence
30
                <220>
                <223>
                                961c-ORF46.1
                <400>
                atggccacaa acgacgacga tgttaaaaaa gctgccactg tggccattgc tgctgcctac
                                                                                                                                                                60
                                                                                                                                                                120
                aacaatggcc aagaaatcaa cggtttcaaa gctggagaga ccatctacga cattgatgaa
35
               gacggcacaa ttaccaaaaa agacgcaact gcagccgatg ttgaagccga cgactttaaa ggtctgggtc tgaaaaaagt cgtgactaac ctgaccaaaa ccgtcaatga aaacaaacaa
                                                                                                                                                                180
                                                                                                                                                                240
               aacgtcgatg ccaaagtaaa agctgcagaa tctgaaatag aaaagttaac aaccaagtta gcagacactg atgccgctt agcagatact gatgccgctc tggatgcaac caccaacgcc ttgaataaat tgggagaaaa tataacgaca tttgctgaag agactaagac aaatatcgta
                                                                                                                                                                300
                                                                                                                                                               360
420
480
               aaaattgatg aaaaattaga agccgtggct gataccgtcg acaagcatgc cgaagcattc aacgatatcg ccgattcatt ggatgaaacc aacactaagg cagacgaagc cgtcaaaacc gccaatgaag ccaaacagac ggccgaagaa accaaacaaa acgtcgatgc caaagtaaaa
                                                                                                                                                                540
40
                                                                                                                                                                600
               gctgcagaaa ctgcagcagg caaagccgaa gctgccgctg gcacagctaa tactgcagcc
gacaaggccg aagctgtcgc tgcaaaagtt accgacatca aagctgatat cgctacgaac
                                                                                                                                                                660
                                                                                                                                                                720
               adagataata ttgctaaaaa agcaaacagt gccgacgtgt acaccagaga agagtctgac agcaaatttg tcagaattga tggtctgaac gctactaccg aaaaattgga cacacgcttg gcttctgctg aaaaatccat tgccgatcac gatactcgcc tgaacggttt ggataaaaca gtgtcagacc tgcgcaaaga aacccgccaa ggccttgcag aacaagccgc gctctccggt ctgttccaac cttacaacgt gggtggatcc ggaggaggag gatcagattt ggcaaacgat tcttttatcc ggcaggttct cgaccgtcag catttcgaac ccgacgggaa ataccaccta
                                                                                                                                                                780
                                                                                                                                                                840
                                                                                                                                                                900
45
                                                                                                                                                                960
                                                                                                                                                                1020
                                                                                                                                                                1080
               ttcggcagca ggggggaact tgccgagcgc agcggccata tcggattggg aaaaatacaa agccatcagt tgggcaacct gatgattcaa caggcggcca ttaaaggaaa tatcggctac
                                                                                                                                                                1140
                                                                                                                                                                1200
               agccatcagt tgggcaacct gatgattcaa caggcggcca ttaaaaggaaa tatcggctac attgtccgct tttccgatca cgggcacgaa gtccattcc ccttcgacaa ccatgcctca cattccgatt ctgatgaagc cggtagtccc gttgacggat ttagccttta ccgcatccat tgggacggat acgaacacca tcccgccgac ggctatgacg ggccacaggg cggcggctat cccgctcca aaggcgcgag ggatatatac agctacgaca taaaaggcgt tgcccaaaat atccgcctca acctgaccga caaccgcagc accggacaac ggcttgccga ccgttccac aatgccggta gtatgctgac gcaaggagta ggcgacggat tcaaacggcc cacccgatac agccccgagc tggacagatc ggcaggagaa attgccggc caggcgatgc cgtgcaggacgatac ggtaaaaaca tcatcggcgc ggcaggagaa attgccggc caggcgatgc cgtgcaggacgatagc gtcaaaaaca gctcaaacat tgctgtcata cacgcctag gtctgcttc caccaaacac
                                                                                                                                                                1260
50
                                                                                                                                                                1320
                                                                                                                                                                1380
                                                                                                                                                                1440
                                                                                                                                                                1500
                                                                                                                                                                1560
                                                                                                                                                                1620
55
                                                                                                                                                               1680
                ataagcgaag gctcaaacat tgctgtcatg cacggcttgg gtctgctttc caccgaaaac
                                                                                                                                                                1740
```

5	gccato aatato ttgggo aaaggo ccttao acctco	cgcg ttta ggca aaat catt tcaa acag	attg tggc tcac ccgc cccg ccgt gcgt	ggca agcca ggca cgtca aaata gccg accg	gt coat coat coat coat coat coat coat coa	caāaa ccca ccta gacaa cgtto tcaaa gacg	accco tcaaa tcaaa attti caaaa acggo gtaaa	a aad a ggg g cgg t gcc t ttg c aad a ggg	tgccq gattq gtcgatq ggagq aaatq gttt	gcac ggag caga gcgg cagc gtca	tgg cata gtta aac	gcata ttcg gcgc acgc acgg tggc	aga ggg gat caa caa aga	agcci aaaa cgca atacc agaa ccaa	gcagca gtcagc tacggc ttgccg ccgtcc aacatc cgccac gtgaaa	1800 1860 1920 1980 2040 2100 2160 2220 2256
10	<210> <211> <212> <213>	153 751 PR3 Ar1	L	ial s	Seque	ence										
15	<220> <223>	961	961c-ORF46.1													
	<400> Met A 1	153 a Thr		Asp 5	Asp	Asp	٧a٦	Lys	Lys 10	Ala	Ala	Thr	val	Ala 15	Ile	
20	Ala A	a Ala	туг 20	Asn	Asn	Gly	Gln	Glu 25	ıle	Asn	Gly	Phe	Lys 30	Ala	Gly	
	Glu Th	ır Ile 35	Yyr	Asp	Ile	Asp	Glu 40	Asp	Gly	Thr	Ile	Thr 45	Lys	Lys	Asp	
25	Ala Th		ı Ala	Asp	٧a٦	Glu 55	Аlа	Asp	Asp	Phe	Lys 60	Gly	Leu	Gly	Leu	
	Lys Ly 65	's Va	Val	Thr	Asn 70	Leu	Thr	Lys	Thr	Va1 75	Asn	Glu	Asn	Lys	G]n 80	
30	Asn Va	l Asp	Ala	Lys 85	٧a٦	Lys	Ala	Ala	Glu 90	Ser	Glu	Ile	Glu	Lys 95	Leu	
	Thr Th	ır Lys	Leu 100	Аlа	Asp	Thr	Asp	А]а 105	Ala	Leu	Ala	Asp	Thr 110	Asp	Ala	
35	Ala Le	u Asp 115		Thr	Thr	Asn	Ala 120	Leu	Asn	Lys	Leu	Gly 125	Glu	Asn	Ile	
	Thr Th		e Ala	Glu	Glu	Thr 135	Lys	Thr	Asn	Ile	۷a٦ 140	Lys	Ile	Asp	Glu	
40	Lys Le 145	u Glu	ı Ala	val	Ala 150	Asp	Thr	val	Asp	Lys 155	His	Ala	Glu	Ala	Phe 160	
	Asn As	p Ile	e Ala	Asp 165	Ser	Leu	Asp	Glu	Thr 170	Asn	Thr	Lys	Ala	Asp 175	Glu	
45	Ala Va	l Lys	180	Аlа	Asn	Glu	Ala	Lys 185	Gln	Thr	Аlа	Glu	Glu 190	Thr	Lys	
	Gln As	n Val 195		Ala	Lys	val	Lys 200	Ala	Ala	Glu	Thr	Ala 205	Ala	Gly	Lys	
<i>E</i> 0	Ala Gl 21		ı Ala	Ala	Gly	Thr 215	Ala	Asn	Thr	Ala	Ala 220	Asp	Lys	Ala	Glu	
50	Ala Va 225	l Ala	ı Ala	Lys	va1 230	Thr	Asp	Ile	Lys	Ala 235	Asp	Ile	Ala	Thr	Asn 240	
	Lys As	p Asr	ıle	Ala 245	Lys	Lys	Ala	Asn	Ser 250	Ala	Asp	∨al	Tyr	Thr 255	Arg	
55	Glu Gl	u Ser	Asp 260	Ser	Lys	Phe	val	Arg 265	Ile	Asp	Gly	Leu	Asn 270	Ala	Thr	

Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala 275 280 285 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu 290 295 300 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly 305 310 315 320Leu Phe Gln Pro Tyr Asn Val Gly Gly Ser Gly Gly Gly Ser Asp 325 330 335 10 Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg Gln His Phe 340 345 350Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg Gly Glu Leu Ala 355 360 365 Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln Ser His Gln Leu  $370 \hspace{1cm} 375 \hspace{1cm} 380$ Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly Asn Ile Gly Tyr 385 390 395 400 20 Ile Val Arg Phe Ser Asp His Gly His Glu Val His Ser Pro Phe Asp 405 410 415 Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly Ser Pro Val Asp 420 425 430 Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu His His Pro 435 440 445 Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Gly Tyr Pro Ala Pro Lys 450 455 460 30 Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val Ala Gln Asn 465 470 475 480 Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln Arg Leu Ala 485 490 495 35 Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly Val Gly Asp  $500 \hspace{1cm} 505 \hspace{1cm} 510$ Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp Arg Ser Gly 515 520 525 Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val Lys Asn Ile 530 540 Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp Ala Val Gln Gly 545 550 555 45 Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu Gly Leu Leu 565 570 575 Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala Asp Met Ala 580 590 50 Gln Leu Lys Asp Tyr Ala Ala Ala Ala Ile Arg Asp Trp Ala Val Gln 595 600 605 Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn Ile Phe Met 610 620 Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly Lys Tyr Gly 625 630 635 640

```
Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln Met Gly Ala
645 650 655
            Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn Phe Ala Asp 660 670
            Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg Asn Ile Arg
675 680 685
            Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr Ser Ser Thr 690 695 700
10
            Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp Gln Arg His
705 710 715 720
            Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro Asn Phe Glu
725 730 735
            Lys His Val Lys Tyr Asp Thr Leu Glu His His His His His 740 745 750
            <210>
                         154
            <211>
                         1773
            <212>
                         DNA
                         Artificial Sequence
            <213>
            <220>
                         961c-741
            <223>
25
            <400>
            atggccacaa acgacgacga tgttaaaaaa gctgccactg tggccattgc tgctgcctac
                                                                                                                             60
            aacaatggcc aagaaatcaa cggtttcaaa gctggagaga ccatctacga cattgatgaa gacggcacaa ttaccaaaaa agacgcaact gcagccgatg ttgaagccga cgactttaaa ggtctgggtc tgaaaaaagt cgtgactaac ctgaccaaaa ccgtcaatga aaacaaacaa
                                                                                                                             120
                                                                                                                             180
                                                                                                                             240
            aacgtcgatg ccaaagtaaa agctgcagaa tctgaaatag aaaagttaac aaccaagtta
gcagacactg atgccgcttt agcagatact gatgccgctc tggatgcaac caccaacgcc
ttgaataaat tgggagaaaa tataacgaca tttgctgaag agactaagac aaatatcgta
                                                                                                                             300
30
                                                                                                                             360
           420
                                                                                                                             480
                                                                                                                             540
                                                                                                                             600
                                                                                                                             660
35
                                                                                                                             720
                                                                                                                             780
                                                                                                                             840
                                                                                                                             900
                                                                                                                             960
                                                                                                                             1020
40
                                                                                                                             1080
                                                                                                                             1140
                                                                                                                             1200
           agccgtttcg actttatccg ccaaatcgaa gtggacgggc agctcattac cttggagagt ggagagttcc aagtatacaa acaaagccat tccgccttaa ccgcctttca gaccgagcaa atacaagatt cggagcattc cgggaagatg gttgcgaaac gccagttcag aatcggcgac atagcgggcg aacatacatc ttttgacaag cttcccgaag gcggcagggc gacatatcgc gggacggcgt tcggttcaga cgatgccggc ggaaaactga cctacaccat agatttcgcc gccaagcagg gaaacggcaa aatcgacat ttgaaatcgc cagaactcaa tgtcgacctg
                                                                                                                             1260
                                                                                                                             1320
                                                                                                                             1380
                                                                                                                             1440
45
                                                                                                                             1500
                                                                                                                             1560
            gccgccgccg atatcaagcc ggatggaaaa cgccatgccg tcatcagcgg ttccgtcctt
                                                                                                                             1620
           tacaaccaag ccgagaaagg cagttactcc ctcggtatct ttggcggaaa agcccaggaa gttgccggca gcgcggaagt gaaaaccgta aacggcatac gccatatcgg ccttgccgcc aagcaactcg agcaccacca ccaccacca tga
                                                                                                                             1680
                                                                                                                             1740
                                                                                                                             1773
50
            <210>
                         155
            <211>
                         590
            <212>
                         PRT
            <213>
                         Artificial Sequence
55
            <220>
            <223>
                         961c-741
```

Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly 20 25 30 Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp 35 40 4510 Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu 50 60 Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln 65 75 80 Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu 85 90 95 15 Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala 100 105 110 Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile 115 120 125 20 Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu 130 140 Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe 145 150 160 25 Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu 165 170 175 Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys  $180 \hspace{1cm} 185 \hspace{1cm} 190$ 30 Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys 195 200 205 Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu 210 215 220 35 Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn 225 235 240 Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg 245 250 255 40 Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr 260 270 Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala 275 280 285 45 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu 290 295 300 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly 305 310 315 50 Leu Phe Gln Pro Tyr Asn Val Gly Gly Ser Gly Gly Gly Val Ala 325 330 335 Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala Pro Leu Asp  $340 \hspace{1.5cm} 345 \hspace{1.5cm} 350$ 55 His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln Ser Val Arg

		355		360		365	
5	Lys Asn 370		Leu Lys Le 37	eu Ala Ala 75	Gln Gly Ala 380	Glu Lys Thr	Tyr
	Gly Asn 385	Gly Asp	Ser Leu As 390	sn Thr Gly	Lys Leu Lys 395	Asn Asp Lys	val 400
10	Ser Arg		Phe Ile Ar 405	rg Gln Ile	Glu Val Asp 410	Gly Gln Leu 415	Ile
	Thr Leu	Glu Ser 420	Gly Glu Ph	he Gln Val 425	Tyr Lys Gln	Ser His Ser 430	Ala
45	Leu Thr	Ala Phe 435	Gln Thr Gl	lu Gln Ile 440	Gln Asp Ser	Glu His Ser 445	Gly
15	Lys Met 450	Val Ala		ln Phe Arg 55	Ile Gly Asp 460	Ile Ala Gly	Glu
	His Thr 465	Ser Phe	Asp Lys Le 470	eu Pro Glu	Gly Gly Arg 475	Ala Thr Tyr	Arg 480
20	Gly Thr		Gly Ser As 485	sp Asp Ala	Gly Gly Lys 490	Leu Thr Tyr 495	Thr
	Ile Asp	Phe Ala . 500	Ala Lys Gl	ln Gly Asn 505	Gly Lys Ile	Glu His Leu 510	Lys
25	Ser Pro	Glu Leu . 515	Asn Val As	sp Leu Ala 520	Ala Ala Asp	Ile Lys Pro 525	Asp
	Gly Lys 530	Arg His		le Ser Gly 35	Ser Val Leu 540	Tyr Asn Gln	Ala
30	Glu Lys 545	Gly Ser	Tyr Ser Le 550	eu Gly Ile	Phe Gly Gly 555	Lys Ala Gln	G1u 560
	Val Ala		Ala Glu Va 565	al Lys Thr	Val Asn Gly 570	Ile Arg His 575	Ile
35	Gly Leu	Ala Ala 580	Lys Gln Le	eu Glu His 585	ніs <b>ніs ні</b> s	His His 590	
40	<210> <211> <212> <213>	156 4170 DNA Artifici	al Sequenc	ce			
	<220> <223>	961c-983					
45	aacaatg gacggca	gcc aagaa caa ttacc	ātcāa cģgt aaaaa agad	tttcaaa gct cgcaact gca	ggagaga cca gccgatg ttg	ccattgc tgct tctacga catt aagccga cgac tcaatga aaac	gatgaa 120 tttaaa 180
50	aacgtcg gcagaca ttgaata aaaattg aacgata	atg ccaaa ctg atgcc aat tggga atg aaaaa tcg ccgat	gtaāa aģct gcttt agca gaaaa tata ttaga agcc tcatt ggat	tgcagaa tct agatact gat aacgaca ttt cgtggct gat tgaaacc aac	gaaatag aaa gccgctc tgg gctgaag aga accgtcg aca actaagg cag	agttaac aacca atgcaac cacca ctaagac aaata agcatgc cgaa acgaagc cgtc	aagtta 300 aacgcc 360 atcgta 420 gcattc 480 aaaacc 540
55	gctgcag gacaagg aaagata agcaaat gcttctg	aaa ctgca ccg aagct ata ttgct ttg tcaga ctg aaaaa	gcagg caaa gtcgc tgca aaaaa agca attga tggt tccat tgcc	agccgaa gct aaaagtt acc aaacagt gcc tctgaac gct cgatcac gat	gccgctg gca gacatca aag gacgtgt aca actaccg aaa actcgcc tga	tcgatgc caaa cagctaa tact ctgatat cgct ccagaga agag aattgga caca acggttt ggat aagccgc gctc	gcagcc 660 acgaac 720 tctgac 780 cgcttg 840 aaaaca 900

```
ctgttccaac cttacaacgt gggtggatcc ggcggaggcg gcacttctgc gcccgacttc aatgcaggcg gtaccggtat cggcagcaac agcagagcaa caacagcgaa atcagcagca
                                                                                                                                                                                                                1080
                       gtatcttacg ccggtatcaa gaacgaaatg tgcaaagaca gaagcatgct ctgtgccggt cgggatgacg ttgcggttac agacagggat gccaaaatca atgcccccc cccgaatctg cataccggag actttccaaa cccaaatgac gcatacaaga atttgatcaa cctcaaacct gcaattgaag caggctatac aggacgcggg gtagagggtag gtatcgtcga cacaggcgaa tccgtcggca gcatatcctt tcccgaacctg tatacgagaa aagaacacgg ctataacgaa
                                                                                                                                                                                                                1140
                                                                                                                                                                                                               1200
1260
1320
                     1380
                                                                                                                                                                                                                1440
                                                                                                                                                                                                                1500
                                                                                                                                                                                                                1560
                                                                                                                                                                                                               1620
1680
1740
10
                                                                                                                                                                                                                1800
                                                                                                                                                                                                               1860
1920
                                                                                                                                                                                                                1980
15
                                                                                                                                                                                                               2040
2100
                                                                                                                                                                                                                2160
                                                                                                                                                                                                                2220
2280
                                                                                                                                                                                                               2340
2400
20
                                                                                                                                                                                                                2460
                                                                                                                                                                                                               2520
2580
                                                                                                                                                                                                                2640
                                                                                                                                                                                                               2700
2760
2820
25
                                                                                                                                                                                                                2880
                                                                                                                                                                                                                2940
                                                                                                                                                                                                                3000
                                                                                                                                                                                                                3060
                                                                                                                                                                                                                3120
30
                                                                                                                                                                                                                3180
                                                                                                                                                                                                               3240
                                                                                                                                                                                                               3300
                                                                                                                                                                                                                3360
                       caaacccaac aggacggtgg aacgtgggaa cagggcggtg ttgaaggcaa aatgcgcggc
agtacccaaa ccgtcggcat tgccgcgaaa accggcgaaa atacgacagc agccgcaca
ctgggcatgg gacgcagcac atggagcgaa aacagtgcaa atgcaaaaac cgacagcatt
                                                                                                                                                                                                                3420
                                                                                                                                                                                                                3480
35
                                                                                                                                                                                                                3540
                      ctgggcatgg gacgcagcac atggagcgaa aacagtgcaa atgcaaaaac cgacagcatt agtctgtttg caggcatacg gcacgatgcg ggcgatatcg gctactcaa aggcctgttc tcctacggac gctacaaaaa cagcatcagc cgcagcaccg gtgcggacga acatgcggaa ggcagcgtca acggcaccg gatgcagcgt gatgcaagcg ggcgactgg gcggtgtcaa cgttccgttt gccgcaacgg gagatttgac ggtcgaaggc ggtctgcgct acgacctgct caaacaggat gcattcgcg aaaaaaggcag tgctttgggc tggagcggca acagcctcac tgaaggcacg ctggtcggac tcgcgggtct gaagctgtcg caacccttga gcgataaagc cgtcctgttt gcaacggcgg gcgtggaacg cgacctgaac ggacgcacgca acacggtaac gggcggcttt accggcgcga ctgcagcaac cggcaagacg ggggcacgca atatgccgca cacccgtctg gttgccggc tggggcggaattc ggcaacggct ggaacggct ggcacgttac agctacgccg gttccaaaca gtacggcaac cacagcggac gagtcggcgt aggctaccgg ttcctcgagc accaccacca ccaccactga
                                                                                                                                                                                                                3600
                                                                                                                                                                                                                3660
                                                                                                                                                                                                                3720
                                                                                                                                                                                                                3780
                                                                                                                                                                                                                3840
40
                                                                                                                                                                                                                3900
                                                                                                                                                                                                                3960
                                                                                                                                                                                                               4020
                                                                                                                                                                                                               4080
4140
                                                                                                                                                                                                                4170
45
                       <210>
                       <211>
                                             1389
                       <212>
                                             PRT
                                            Artificial Sequence
                        <213>
                       <220>
50
                       <223>
                                            961c-983
                      Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile
1 5 10
55
                      Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly 20 25 30
```

Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp 40 45Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu 50 60 Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln 65 75 80 Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu 85 90 95 10 Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala
100 105 110 Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile 115 120 125 15 Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu 130 135 140 Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe 145 150 155 160 20 Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu 165 170 175 Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys 180 185 190 25 Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys 195 200 205 Ala Glu Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu 210 215 220 30 Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn 225 230 240 Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg 245 250 255 35 Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr  $260 \hspace{1cm} 265 \hspace{1cm} 270 \hspace{1cm}$ Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala 275 280 285 40 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu 290 295 300 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly 305 310 315 320 45 Leu Phe Gln Pro Tyr Asn Val Gly Gly Ser Gly Gly Gly Gly Thr Ser 325 330 335 Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser Asn Ser Arg 340 345 35050 Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly Ile Lys Asn 355 360 365 Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg Asp Asp Val 370 375 380 Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro Pro Asn Leu 385 390 400

His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys Asn Leu Ile 405 410 415 Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg Gly Val Glu
420 425 430 Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile Ser Phe Pro 435 440 445 Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn Tyr Lys Asn 450 455 460 10 Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly Gly Gly Lys
465 470 475 480 Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu Thr Glu Ala 485 490 495 15 Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His Ile Asp Leu 500 505 510 Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg Pro Ala Gly 515 520 525 20 Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr Asn Asp Glu 530 540 Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala Trp Val Lys 545 550 560 25 Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe Gly Thr Thr 565 570 575 Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn Ser Glu Glu 580 585 590 30 Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp Lys Thr Asp 595 600 605 Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn Leu Ser Tyr 610 620 35 His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr Gly Asn Asp 625 630 635 640 Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe Tyr Glu Lys 645 650 655 40 Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp Arg Ser Gly 660 665 670 Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr Glu Pro Leu 675 680 45 Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp Cys Leu Ser 690 695 700 Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn Pro Ile Gln 705 710 720 50 Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly Thr Ala Ala 725 730 735 Leu Leu Eln Lys Tyr Pro Trp Met Ser Asn Asp Asn Leu Arg Thr 740 745 750 55 Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly Val Asp Ser 765

	Lys	770	GIY	Trp	GIY	Leu	775	Asp	АІа	GIY	Lys	780	Met	Asn	GIY	Pro
5	Ala 785	Ser	Phe	Pro	Phe	Gly 790	Asp	Phe	Thr	ΑΊа	Asp 795	Thr	Lys	Gly	Thr	Ser 800
	Asp	Ile	Ala	Tyr	Ser 805	Phe	Arg	Asn	Asp	11e 810	Ser	Gly	Thr	Gly	Gly 815	Leu
10	Ile	Lys	Lys	Gly 820	Gly	Ser	Gln	Leu	G]n 825	Leu	ніѕ	Gly	Asn	Asn 830	Thr	туг
	Thr	Gly	Lys 835	Thr	Ile	ıle	Glu	Gly 840	Gly	Ser	Leu	val	Leu 845	туг	Gly	Asn
15	Asn	Lys 850	ser	Asp	Met	Arg	va1 855	Glu	Thr	Lys	Gly	Ala 860	Leu	Ile	Tyr	Asn
	Gly 865	ΑΊа	Ala	Ser	Gly	Gly 870	Ser	Leu	Asn	ser	Asp 875	Gly	Ile	val	Tyr	Leu 880
20	Ala	Asp	Thr	Asp	G]n 885	Ser	Gly	Ala	Asn	G]u 890	Thr	val	нis	Ile	Lys 895	Gly
	Ser	Leu	Gln	Leu 900	Asp	Glу	Lys	Gly	Thr 905	Leu	Tyr	Thr	Arg	Leu 910	Gly	Lys
25	Leu	Leu	Lys 915	val	Asp	Gly	Thr	Ala 920	ıle	ıle	Gly	Gly	Lys 925	Leu	Tyr	Met
	Ser	Ala 930	Arg	Gly	Lys	Gly	А]а 935	GТу	Tyr	Leu	Asn	ser 940	Thr	Gly	Arg	Arg
30	Val 945	Pro	Phe	Leu	Ser	А1а 950	ΑΊа	Lys	Ile	Gly	G]n 955	Asp	Tyr	Ser	Phe	Phe 960
	Thr	Asn	Ile	Glu	Thr 965	Asp	Gly	Gly	Leu	Leu 970	Ala	Ser	Leu	Asp	Ser 975	val
<i>35</i>	Glu	Lys	Thr	А]а 980	Gly	Ser	Glu	Gly	Asp 985	Thr	Leu	Ser	Tyr	Туг 990	val	Arg
	Arg	Gly	Asn 995	Ala	Аlа	Arg	Thr	Ala 1000		Ala	Ala	Ala	ніs 1005		ΑΊα	Pro
40	Ala	Gly 1010		Lys	His	Ala	Val 1015		Gln	GТу	Gly	Ser 1020		Leu	Glu	Asn
	Leu 1025		۷a٦	Glu	Leu	Asp 1030	Ala )	Ser	Glu	Ser	Ser 1035		Thr	Pro	Glu	Thr 1040
45	٧a٦	Glu	Thr	Ala	Ala 1045		Asp	Arg	Thr	Asp 1050		Pro	Gly	Ile	Arg 1055	
	Tyr	Gly	Ala	Thr 1060		Arg	Ala	Ala	Ala 1065		٧a٦	Gln	His	А]а 1070		Ala
50	Ala	Asp	Gly 1075		Arg	IJе	Phe	Asn 1080		Leu	Ala	Ala	Thr 1085		Туг	Ala
50	Asp	Ser 1090		Ala	Ala	His	Ala 1095		Met	Gln	Gly	Arg 1100		Leu	Lys	Ala
	va1 1105		Asp	Gly	Leu	Asp 1110	His )	Asn	Glу	Thr	Gly 1115		Arg	∨al	ile	Ala 1120
55	Gln	Thr	Gln	Gln	Asp 1125		Glу	Thr	Trp	Glu 1130		Gly	Gly	val	Glu 1135	

```
Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala Lys Thr Gly 1140 1145 1150
         Glu Asn Thr Thr Ala Ala Ala Thr Leu Gly Met Gly Arg Ser Thr Trp
1155 1160 1165
         Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser Leu Phe Ala
1170 1175 1180
         Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys Gly Leu Phe 1185 1190 1195 120
10
         Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr Gly Ala Asp
1205 1210 1215
         Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln Leu Gly Ala
1220 1225 1230
15
         Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp Leu Thr Val
1235 1240 1245
         Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys Gln Asp Ala Phe Ala Glu
1250 1255 1260
20
         Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser Leu Thr Glu Gly Thr
1265 1270 1275 128
         Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln Pro Leu Ser Asp Lys
1285 1290 1295
25
         Ala Val Leu Phe Ala Thr Ala Gly Val Glu Arg Asp Leu Asn Gly Arg
1300 1305 1310
         Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly Ala Thr Ala Ala Thr Gly 1315 1320 1325
30
         Lys Thr Gly Ala Arg Asn Met Pro His Thr Arg Leu Val Ala Gly Leu
         Gly Ala Asp Val Glu Phe Gly Asn Gly Trp Asn Gly Leu Ala Arg Tyr
1345 1350 1355 136
35
         Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn His Ser Gly Arg Val Gly 1365 1370 1375
         Val Gly Tyr Arg Phe Leu Glu His His His His His
                                                1385
40
         <210>
                   158
         <211>
                   2304
         <212>
                   DNA
         <213>
                   Artificial Sequence
         <220>
45
                   961cL-ORF46.1
         <223>
         <400>
                   158
         atgaaacact ttccatccaa agtactgacc acagccatcc ttgccacttt ctgtagcggc
                                                                                             60
         gcactggcag ccacaaacga cgacgatgtt aaaaaagctg ccactgtggc cattgctgct
                                                                                             120
                                                                                             180
         gcctacaaca atggccaaga aatcaacggt ttcaaagctg gagagaccat ctacgacatt
50
         gatgaagacg gcacaattac caaaaaagac gcaactgcag ccgatgttga agccgacgac
tttaaaggtc tgggtctgaa aaaagtcgtg actaacctga ccaaaaccgt caatgaaaac
                                                                                             240
                                                                                             300
         aaacaaaacg tcgatgccaa agtaaaagct gcagaatctg aaatagaaaa gttaacaacc
                                                                                             360
         aagttagcag acactgatgc cgctttagca gatactgatg ccgctctgga tgcaaccacc aacgccttga ataaattggg agaaaatata acgacatttg ctgaagagac taagacaaat
                                                                                             420
                                                                                             480
         atcotaaaaa ttgatgaaaa attagaagcc gtggctgata ccotcogacaa gcatgccgaa
                                                                                             540
55
         gcattcaacg atatcgccga ttcattggat gaaaccaaca ctaaggcaga cgaagccgtc
                                                                                             600
         aaaaccgcca atgaagccaa acagacggcc gaagaaacca aacaaaacgt cgatgccaaa
                                                                                             660
```

720

gtaaaagctg cagaaactgc agcaggcaaa gccgaagctg ccgctggcac agctaatact

```
gcagccgaca aggccgaagc tgtcgctgca aaagttaccg acatcaaagc tgatatcgct
                                                                                                                                                     780
                acgaacaaag ataatattgc taaaaaagca aacagtgccg acgtgtacac cagagaagag
tctgacagca aatttgtcag aattgatggt ctgaacgcta ctaccgaaaa attggacaca
                                                                                                                                                     840
                                                                                                                                                     900
                cgcttggctt ctgctgaaaa atccattgcc gatcacgata ctcgcctgaa cggtttggat
                                                                                                                                                     960
                aaaacagtgt cagacctgcg caaagaaacc cgccaaggcc ttgcagaaca agccgcgctc
tccggtctgt tccaacctta caacgtgggt ggatccggag gaggaggatc agatttggca
aacgattctt ttatccggca ggttctcgac cgtcagcatt tcgaacccga cgggaaatac
                                                                                                                                                     1020
                                                                                                                                                     1080
                                                                                                                                                     1140
                aacgattett ttateegga ggttetegae egteageatt tegaaceega egggaaatae cacetatteg geageaggg ggaacttgee gagegeageg gecatategg attegggaaaa atacaaagee ateagttggg caacetgatg atteaacagg eggeeattaa aggaaaatae ggetacattg teegettte egateaeggg cacgaagtee atteeecett egacaaceat geeteaatt eegateega tgaageegg agteeeggt agteeegtt aeggattag eetttaeege atecattggg aeggataega acaceatee geegaegget atgaegggee acagggegge ggetateegg eteceaaagg egegagggat atatacaget aegacataaa aggeegttgee eaaaatatee geeteaacet gaeegaeaae egeageaegg aeggatteaa aegeegeegg teegataeagee eegataeegg eagateega aatgeegeeg aatgeegeeg aageetteaa eggeeaetgeagatategtta aaaacateat eggeeggga ggaaattg teggeegaag egateegta
                                                                                                                                                     1200
                                                                                                                                                     1260
10
                                                                                                                                                     1320
                                                                                                                                                     1380
                                                                                                                                                     1440
                                                                                                                                                     1500
                                                                                                                                                     1560
                                                                                                                                                     1620
15
                                                                                                                                                     1680
                                                                                                                                                     1740
                gatatcgita aaaacatcat cggcgcggca ggagaaattg tcggcgcagg cgatgccgtg
                cagggcataa gcgaaggctc aaacattgct gtcatgcacg gcttgggtct gctttccacc gaaaacaaga tggcgcgcat caacgatttg gcagatatgg cgcaactcaa agactatgcc
                                                                                                                                                     1800
                                                                                                                                                     1860
                gcagcagcca tccgcgattg ggcagtccaa aaccccaatg ccgcacaagg catagaagcc
gtcagcaata tctttatggc agccatcccc atcaaaggga ttggagctgt tcggggaaaa
                                                                                                                                                     1920
                                                                                                                                                     1980
                tacggcttgg gcggcatcac ggcacatcct atcadaggga ttggagctgt tcggggadad tacggcttgg gcggcatcac ggcacatcct atcadagcggt cgcagatggg cgcgatcgca ttgccgaaag ggaaatccgc cgtcagcgac aattttgccg atgcggcata cgccaaatac ccgtccctt accattccg aaatatccgt tcadaacttgg agcagcgtta cggcaaagaa aacatcacct cctcaaccgt gccgccgtca aacggcaaaa atgtcaaact ggcagaccaa cgccacccga agacaggcgt accgtttgac ggtaaagggt ttccgaattt tgagaagcac
                                                                                                                                                     2040
20
                                                                                                                                                     2100
                                                                                                                                                     2160
                                                                                                                                                     2220
                                                                                                                                                     2280
                gtgaaatatg atacgtaact cgag
                                                                                                                                                     2304
25
                                159
765
                <210>
                 <211>
                 <212>
                 <213>
                                Artificial Sequence
                 <220>
                <223>
                                961cL-ORF46.1
30
                Met Lys His Phe Pro Ser Lys Val Leu Thr Thr Ala Ile Leu Ala Thr
1 5 10 15
                Phe Cys Ser Gly Ala Leu Ala Ala Thr Asn Asp Asp Asp Val Lys Lys 20 25 30
35
                Ala Ala Thr Val Ala Ile Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile
35 40 45
                Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly 50 60
40
                Thr Ile Thr Lys Lys Asp Ala Thr Ala Ala Asp Val Glu Ala Asp Asp 65 70 75 80
                Phe Lys Gly Leu Gly Leu Lys Lys Val Val Thr Asn Leu Thr Lys Thr 85 90 95
45
                Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu
100 105
                Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala 115 120 125
50
                Leu Ala Asp Thr Asp Ala Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn 130 140
                Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn 145 150 160
55
                Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala Asp Thr Val Asp
```

						165					170					175	
į	5	Lys	His	Ala	Glu 180	Ala	Phe	Asn	Asp	11e 185	Ala	Asp	Ser	Leu	Asp 190	Glu	Thr
		Asn	Thr	Lys 195	Ala	Asp	Glu	Ala	va1 200	Lys	Thr	Ala	Asn	G1u 205	Ala	Lys	Gln
1	0	Thr	Ala 210	Glu	Glu	Thr	Lys	Gln 215	Asn	۷a٦	Asp	Ala	Lys 220	۷al	Lys	Ala	Ala
		Glu 225	Thr	Ala	Ala	Gly	Lys 230	Ala	Glu	Ala	Ala	Ala 235	Gly	Thr	Ala	Asn	Thr 240
1	5	Ala	Ala	Asp	Lys	Ala 245	Glu	Ala	val	Ala	Ala 250	Lys	Va1	Thr	Asp	11e 255	Lys
		Ala	Asp	Ile	Ala 260	Thr	Asn	Lys	Asp	Asn 265	Ile	Ala	Lys	Lys	Ala 270	Asn	ser
9	0	Ala	Asp	Va 1 275	Tyr	Thr	Arg	Glu	Glu 280	Ser	Asp	Ser	Lys	Phe 285	Val	Arg	Ile
۷	o .	Asp	Gly 290	Leu	Asn	Ala	Thr	Thr 295	Glu	Lys	Leu	Asp	Thr 300	Arg	Leu	Ala	Ser
		Ala 305	Glu	Lys	Ser	Ile	Ala 310	Asp	His	Asp	Thr	Arg 315	Leu	Asn	Gly	Leu	Asp 320
2	25					325		Arg			330					335	
		Gln	Ala	Ala	Leu 340	Ser	Gly	Leu	Phe	Gln 345	Pro	Tyr	Asn	∨al	Gly 350	Gly	Ser
3	80	Gly	Gly	G]y 355	Gly	ser	Asp	Leu	Ala 360	Asn	Asp	Ser	Phe	11e 365	Arg	Gln	val
		Leu	Asp 370	Arg	Gln	His	Phe	G1u 375	Pro	Asp	Gly	Lys	Tyr 380	ніѕ	Leu	Phe	Gly
3	25	385					390	Glu	_			395					400
						405		Gly			410				_	415	
4	0				420			Ile		425					430		_
				435			-	Asn	440					445			
4	5		450					Gly 455					460				
		465					470	Ala	_			475					480
5	50					485		Gly			490					495	
			-		500			Ile		505					510		
5	55			515				Asp	520					525			
		Thr	Gln	Gly	٧a٦	Gly	Asp	Gly	Phe	Lys	Arg	Ala	Thr	Arg	Tyr	ser	Pro

	530		535	540
5	Glu Leu 545	Asp Arg Ser Gly 550	Asn Ala Ala Glu Ala 555	Phe Asn Gly Thr Ala 560
	Asp Ile	Val Lys Asn Ile 565	Ile Gly Ala Ala Gly 570	Glu Ile Val Gly Ala 575
10	Gly Asp	Ala Val Gln Gly 580	Ile Ser Glu Gly Ser 585	Asn Ile Ala Val Met 590
10	His Gly	Leu Gly Leu Leu 595	Ser Thr Glu Asn Lys 600	Met Ala Arg Ile Asn 605
	Asp Leu 610		Gln Leu Lys Asp Tyr 615	Ala Ala Ala Ile 620
15	Arg Asp 625	Trp Ala Val Gln 630	Asn Pro Asn Ala Ala 635	Gln Gly Ile Glu Ala 640
	Val Ser	Asn Ile Phe Met 645	Ala Ala Ile Pro Ile 650	Lys Gly Ile Gly Ala 655
20	Val Arg	Gly Lys Tyr Gly 660	Leu Gly Gly Ile Thr 665	Ala His Pro Ile Lys 670
	Arg Ser	Gln Met Gly Ala 675	Ile Ala Leu Pro Lys 680	Gly Lys Ser Ala Val 685
25	Ser Asp 690		Ala Ala Tyr Ala Lys 695	Tyr Pro Ser Pro Tyr 700
	His Ser 705	Arg Asn Ile Arg 710	Ser Asn Leu Glu Gln 715	Arg Tyr Gly Lys Glu 720
30	Asn Ile	Thr Ser Ser Thr 725	Val Pro Pro Ser Asn 730	Gly Lys Asn Val Lys 735
	Leu Ala	Asp Gln Arg His 740	Pro Lys Thr Gly Val 745	Pro Phe Asp Gly Lys 750
35	Gly Phe	Pro Asn Phe Glu 755	Lys His Val Lys Tyr 760	Asp Thr 765
40	<210> <211> <212> <213>	160 1839 DNA Artificial Seque	ence	
40	<220> <223>	961cL-741		
45				ttgccacttt ctgtagcggc 60 ccactgtggc cattgctgct 120
	gatgaag tttaaag aaacaaa	acg gcacaattac ca gtc tgggtctgaa aa acg tcgatgccaa ag	aaaaagac gcaactgcag aagtcgtg actaacctga taaaagct gcagaatctg	gagagaccat ctacgacatt 180 ccgatgttga agccgacgac 240 ccaaaaccgt caatgaaaac 300 aaatagaaaa gttaacaacc 360
50	aacgcct atcgtaa gcattca	tga ataaattggg ag aaa ttgatgaaaa at acg atatcgccga tt	aaaatata acgacatttg tagaagcc gtggctgata cattggat gaaaccaaca	ccgctctgga tgcaaccacc 420 ctgaagagac taagacaaat 480 ccgtcgacaa gcatgccgaa 540 ctaaggcaga cgaagccgtc 600
55	gtaaaag gcagccg acgaaca tctgaca	ctg cagaaactgc ag aca aggccgaagc tg aag ataatattgc ta gca aatttgtcag aa	caggcaaa gccgaagctg tcgctgca aaagttaccg aaaaagca aacagtgccg ttgatggt ctgaacgcta	aacaaaacgt cgatgccaaa 660 ccgctggcac agctaatact 720 acatcaaagc tgatatcgct 780 acgtgtacac cagagaagag 840 ctaccgaaaa attggacaca 900 ctcgcctgaa cggtttggat 960

1020

aaaacagtgt cagacctgcg caaagaaacc cgccaaggcc ttgcagaaca agccgcgctc

```
tccggtčtgt tccaaccita caacgtgggt ggatccggag ggggtggtgt cgccgccgac
                                                                                                             1080
            atcggtgcgg ggcttgccga tgcactaacc gcaccgctcg accataaaga caaaggtttg
cagtctttga cgctggatca gtccgtcagg aaaaacgaga aactgaagct ggcggcacaa
                                                                                                             1140
                                                                                                             1200
            ggtgcggaaa aaacttatgg aaacggtgac agcctcaata cgggcaaatt gaagaacgac
aaggtcagcc gtttcgactt tatccgccaa atcgaagtgg acgggcagct cattaccttg
                                                                                                             1260
                                                                                                             1320
            gagagtggag agttccaagt atacaaacaa agccattccg ccttaaccgc ctttcagacc
gagcaaatac aagattcgga gcattccggg aagatggttg cgaaacgcca gttcagaatc
ggcgacatag cgggcgaaca tacatctttt gacaagcttc ccgaaggcgg cagggcgaca
                                                                                                             1380
                                                                                                             1440
                                                                                                             1500
            tatcgcggga cggcgttcgg ttcagacgat gccggcggaa aactgaccta caccatagat
ttcgccgcca agcagggaaa cggcaaaatc gaacatttga aatcgccaga actcaatgtc
                                                                                                             1560
10
                                                                                                             1620
            gacctggccg ccgccgatat caagccggat ggaaaacgcc atgccgtcat cagcggttcc gtcctttaca accaagccga gaaaggcagt tactccctcg gtatctttgg cggaaaagcc caggaagttg ccggcagcgc ggaagtgaaa accgtaaacg gcatacgcca tatcggcctt
                                                                                                             1680
                                                                                                             1740
                                                                                                             1800
            gccgccaagc aactcgagca ccaccaccac caccactga
                                                                                                             1839
15
            <210>
                       612
            <211>
            <212>
                       PRT
                       Artificial Sequence
            <220>
<223>
                       961cL-741
20
            Met Lys His Phe Pro Ser Lys Val Leu Thr Thr Ala Ile Leu Ala Thr
1 5 10 15
            Phe Cys Ser Gly Ala Leu Ala Ala Thr Asn Asp Asp Val Lys Lys 20 25 30
25
            Ala Ala Thr Val Ala Ile Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile
35 40
            Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly 50 55
30
            Thr Ile Thr Lys Lys Asp Ala Thr Ala Ala Asp Val Glu Ala Asp Asp 65 70 75 80
            Phe Lys Gly Leu Gly Leu Lys Lys Val Val Thr Asn Leu Thr Lys Thr 85 90 95
35
            Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu
100 105
            Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala
115 120 125
40
            Leu Ala Asp Thr Asp Ala Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn 130 140
            Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn 145 155 160
45
            Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala Asp Thr Val Asp
165 170 175
            Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr
180 185
50
           Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala Asn Glu Ala Lys Gln
195 200 205
            Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala 210 215 220
55
            Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Gly Thr Ala Asn Thr
225 230 240
```

Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys Val Thr Asp Ile Lys 245 250 255 Ala Asp Ile Ala Thr Asn Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser 260 270 Ala Asp Val Tyr Thr Arg Glu Glu Ser Asp Ser Lys Phe Val Arg Ile 275 280 285 Asp Gly Leu Asn Ala Thr Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser 290 295 300 10 Ala Glu Lys Ser Ile Ala Asp His Asp Thr Arg Leu Asn Gly Leu Asp 305 310 320 Lys Thr Val Ser Asp Leu Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu 325 330 335 15 Gln Ala Ala Leu Ser Gly Leu Phe Gln Pro Tyr Asn Val Gly Gly Ser 340 345 350Gly Gly Gly Val Ala Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala 355  $\phantom{-}360\phantom{0}$ 20 Leu Thr Ala Pro Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr 370 380 Leu Asp Gln Ser Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln 385 390 395 25 Gly Ala Glu Lys Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys 405 410 415Leu Lys Asn Asp Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu
420 425 430 30 Val Asp Gly Gln Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr 435 440 445 Lys Gln Ser His Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln 450 460 35 Asp Ser Glu His Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile 465 470 480 Gly Asp Ile Ala Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly 485 490 495 40 Gly Arg Ala Thr Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly 500 505 510 Gly Lys Leu Thr Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly 515 520 525 45 Lys Ile Glu His Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala 530 535 540 Ala Asp Ile Lys Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser 545 550 555 560 50 Val Leu Tyr Asn Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe 565 570 575 Gly Gly Lys Ala Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val 580 585 590 55 Asn Gly Ile Arg His Ile Gly Leu Ala Ala Lys Gln Leu Glu His His 595 600 605

```
His His His His
                   <210>
                                      162
                                      4218
                   <211>
                   <212>
                                      DNA
                   <213>
                                      Artificial Sequence
                   <220>
                   <223>
                                      961cL-983
10
                   <400>
                                      162
                  atgaaacact ttccatccaa agtactgacc acagccatcc ttgccacttt ctgtagcggc
                                                                                                                                                                                         60
                  gcactggcag ccacaaacga cgacgatgtt aaaaaagctg ccactgtggc cattgctgct
                                                                                                                                                                                         120
                  gcctacaaca atggccaaga aatcaacggt ttcaaagctg gagagaccat ctacgacatt gatgaagacg gcacaattac caaaaaagac gcaactgcag ccgatgttga agccgacgac tttaaaggtc tgggtctgaa aaaagtcgtg actaacctga ccaaaaccgt caatgaaaac
                                                                                                                                                                                         180
                                                                                                                                                                                         240
                                                                                                                                                                                          300
                  aaacaaaacg tcgatgccaa agtaaaagct gcagaatctg aaatagaaaa gttaacaacc aagttagcag acactgatgc cgctttagca gatactgatg ccgctctgga tgcaaccacc aacgccttga ataaattggg agaaaatata acgacatttg ctgaagagac taagacaaat
                                                                                                                                                                                          360
                                                                                                                                                                                          420
                                                                                                                                                                                         480
                                                                                                                                                                                         540
                  atcgtaaaaa ttgatgaaaa attagaagcc gtggctgata ccgtcgacaa gcatgccgaa
                  gcattcaacg atatcgccga ttcattggat gaaaccaaca ctaaggcaga cgaagccgtc aaaaccgcca atgaagccaa acagacggcc gaagaaacca aacaaaacgt cgatgccaaa
                                                                                                                                                                                         600
                                                                                                                                                                                         660
20
                  gtaaaagctg cagaaactgc agcaggcaaa gccgaagctg ccgctggcac agctaatact gcagccgaca aggccgaagc tgtcgctgca aaagttaccg acatcaaagc tgatatcgct acgaacaaag ataatattgc taaaaaagca aacagtgccg acgtgtacac cagagaagag tctgacagca aattgatggt ctgaacgcta ctaccgaaaa attggacaca
                                                                                                                                                                                         720
                                                                                                                                                                                         780
                                                                                                                                                                                         840
                                                                                                                                                                                         900
                  cgcttggctt ctgctgaaaa atccattgcc gatcacgata ctcgcctgaa cggtttggat aaaacagtgt cagacctgcg caaagaaacc cgccaaggcc ttgcagaaca agccgcgctctccggtctgt tccaacctta caacgtgggt ggatccggcg gaggcggcac ttctgcgccc
                                                                                                                                                                                         960
                                                                                                                                                                                         1020
25
                                                                                                                                                                                         1080
                  gacttcaatg caggcggtac cggtatcggc agcaacagca gagcaacaac agcgaaatca gcagcagtat cttacgccgg tatcaagaac gaaatgtgca aagacagaag catgctctgt gccggtcggg atgacgttgc ggttacagac agggatgcca aaatcaatgc ccccccccg
                                                                                                                                                                                         1140
                                                                                                                                                                                         1200
                                                                                                                                                                                         1260
                  aatctgcata ccggagactt tccaaaccca aatgacgcat acaagaattt gatcaacctc
                                                                                                                                                                                         1320
                  aaacctgcaa ttgaagcagg ctatacagga cgcggggtag aggtaggtat cgtcgacaca ggcgaatccg tcggcagcat atcetttccc gaactgtatg gcagaaaaga acacggctat aacgaaaatt acaaaaacta tacggcgtat atgcggaagg aagcgcctga agacggaggc ggtaaagaca ttgaagcttc tttcgacgat gaggccgtta tagagactga agcaaagccg acggatatcc gccacgtaaa agaaatcgga cacatcgatt tggtctccca tattattggc
                                                                                                                                                                                         1380
30
                                                                                                                                                                                         1440
                                                                                                                                                                                         1500
                                                                                                                                                                                         1560
                                                                                                                                                                                         1620
                  gggcgttccg tggacggcag acctgcaggc ggtattgcgc ccgatgcgac gctacacata
                                                                                                                                                                                         1680
                  atgaatacga atgatgaaac caagaacgaa atgatggttg cagccatccg caatgcattgg gtcaagctgg gcgaacgtgg cgtgcgcatc gtcaataaca gttttggaac aacatcgagg gcaggcactg ccgacctttt ccaaatagcc aattcggagg agcagtaccg ccaagcgttg ctcgactatt ccggcggtga taaaacagac gagggtatcc gcctgatgca acagagcgat tacggcaacc tgtcctacca catccgtaat aaaaacatgc ttttcatctt ttcgacaggc aatgacgcac aagctcagcc caacacatat gccctattgc cattttatga aaaagacgct
                                                                                                                                                                                         1740
35
                                                                                                                                                                                         1800
                                                                                                                                                                                         1860
                                                                                                                                                                                         1920
                                                                                                                                                                                         1980
                                                                                                                                                                                         2040
                  caaaaaggca ttatcacagt cgcaggcgta gaccgcagtg gagaaaagtt caaacgggaa atgtatggag aaccgggtac agaaccgctt gagtatggct ccaaccattg cggaattact gccatgtggt gcctgtcggc accctatgaa gcaagcgtcc gtttcacccg tacaaacccg
                                                                                                                                                                                         2100
40
                                                                                                                                                                                         2160
                                                                                                                                                                                         2220
                  attcaaattg ccggaacatc cttttccgca cccatcgtaa ccggcacggc ggctctgctg
                                                                                                                                                                                         2280
                  ctgcagaaat accogtggat gagcaacgac aacctgcgta ccacgttgct gacgacggct caggacatcg gtgcagtcgg cgtggacagc aagttcggct ggggactgct ggatgcgggt aaggccatga acggacgcg gtcctttcg ttcggcggtt ttaccgccga tacgaaaggt
                                                                                                                                                                                         2340
                                                                                                                                                                                         2400
                                                                                                                                                                                         2460
                 aaggccatga acggacccgc gtcctttccg ttcggcgact ttaccgccga tacgaaaggt acatccgata ttgcctactc cttccgtaac gacatttcag gcacgggcgg cctgatcaaa aaaggcggca gccaactgca actgcacggc aacaacacct atacgggcaa aaccattatc gaaggcggtt cgctggttt gtacggcaac aacaaatcgg atatgcgcgt cgaaaccaaa ggtgcgctga tttataacgg ggcggcatcc ggcggcagcc tgaaccagcag cggcattgtc tatctggcag ataccgacca atccggcgca aacgaaaccg tacacatcaa aggcagtctg cagctggacg gcaaaggtac gctgtacaca cgtttgggca aactgctgaa aggcagtctg acggcgatta tcggcggcaa gctgtacatg tcggcacgcg gcaagggggc aggctatctc aacagtaccg gacgacgtgt tcccttcctg agtgccgca aaatcgggca ggattattct tcttcacaa acatcgaaac cgacggcggc ctgctggct tccctcgacag cgtcgaaaa acagcgggaa gtgaaggcga cacgctgtcc tattatgtcc gtcgcgcaa tgcggcacg
                                                                                                                                                                                         2520
45
                                                                                                                                                                                         2580
                                                                                                                                                                                         2640
                                                                                                                                                                                         2700
                                                                                                                                                                                         2760
                                                                                                                                                                                         2820
                                                                                                                                                                                         2880
50
                                                                                                                                                                                         2940
                                                                                                                                                                                         3000
                  acagegggca gtgaaggega caegetgtee tattatgtee gtegeggeaa tgeggeacgg
                                                                                                                                                                                         3060
                  actgcttcgg cagcggcaca ttccgcgccc gccggtctga aacacgccgt agaacagggc
ggcagcaatc tggaaaacct gatggtcgaa ctggatgcct ccgaatcatc cgcaacaccc
                                                                                                                                                                                         3120
                                                                                                                                                                                         3180
                  gagacqqttq aaactqcqqc aqccqaccgc acagatatgc cqgqcatccq cccctacggc
                                                                                                                                                                                         3240
                  gcaactttcc gcgcagcggc agccgtacag catgcgaatg ccgccgacgg tgtacgcatc ttcaacagtc tcgccgctac cgtctatgcc gacagtaccg ccgcccatgc cgatatgcag
55
                                                                                                                                                                                         3300
                                                                                                                                                                                         3360
```

3420

```
ggacgccgcc tgaaagccgt atcggacggg ttggaccaca acggcacggg tctgcgcgtc atcgcgcaaa cccaacagga cggtggaacg tgggaacagg gcggtgttga aggcaaaatg cgcggcagta cccaaaccgt cggcattgcc gcgaaaaccg gcgaaaatac gacagcagcc
                                                                                                        3480
                                                                                                        3540
          gccacactgg gcatgggacg cagcacatgg agcgaaaaca gtgcaaatgc aaaaaccgac agcattagtc tgtttgcagg catacggcac gatgcgggcg atatcggcta tctcaaaggc ctgttctcct acggacgcta caaaaacagc atcagccgca gcaccggtgc ggacgaacat
                                                                                                        3600
                                                                                                        3660
                                                                                                        3720
          gcggaaggca gcgtcaacgg cacgctgatg cagctgggcg cactgggcgg tgtcaacgtt
                                                                                                        3780
          ccgtttgccg caacgggaga tttgacggtc gaaggcggtc tgcgctacga cctgctcaaa caggatgcat tcgccgaaaa aggcagtgct ttgggctgga gcggcaacag cctcactgaa
                                                                                                        3840
                                                                                                        3900
           ggcacgctgg tcggactcgc gggtctgaag ctgtcgcaac ccttgagcga taaagccgtc
                                                                                                        3960
          ctgtttgcaa cggcgggcgt ggaacgcgac ctgaacggac gcgactacac ggtaacgggc
ggctttaccg gcgcgactgc agcaaccggc aagacggggg cacgcaatat gccgcacacc
10
                                                                                                        4020
                                                                                                        4080
           cgtctggttg ccggcctggg cgcggatgtc gaattcggca acggctggaa cggcttggca
                                                                                                        4140
                                                                                                        4200
          cgttacaget acgceggite caaacagtae ggcaaccaca geggaegagt eggegtagge
          taccggttct gactcgag
                                                                                                        4218
           <210>
                     163
15
                     1403
           <211>
           <212>
                     PRT
                     Artificial Sequence
           <213>
           <220>
                     961cL-983
           <223>
20
          Met Lys His Phe Pro Ser Lys Val Leu Thr Thr Ala Ile Leu Ala Thr
1 10 15
          Phe Cys Ser Gly Ala Leu Ala Ala Thr Asn Asp Asp Val Lys Lys 20 25 30
25
          Ala Ala Thr Val Ala Ile Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile
35 40 45
          Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly 50 60
30
          Thr Ile Thr Lys Lys Asp Ala Thr Ala Ala Asp Val Glu Ala Asp Asp 65 70 75 80
          Phe Lys Gly Leu Gly Leu Lys Lys Val Val Thr Asn Leu Thr Lys Thr 85 90 95
35
          Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu
100 105 110
          Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala
115 120 125
40
          Leu Ala Asp Thr Asp Ala Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn 130 140
          Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn 145 155 160
45
          Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala Asp Thr Val Asp
165 170 175
          Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr 180 \hspace{1cm} 185
50
          Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala Asn Glu Ala Lys Gln
195 200 205
          Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala
210 215 220
55
          Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Gly Thr Ala Asn Thr
225 230 240
```

	Ala	Ala	Asp	Lys	Ala 245	Glu	Ala	٧a٦	Ala	А]а 250	Lys	٧al	Thr	Asp	11e 255	Lys
5	Ala	Asp	Ile	Ala 260	Thr	Asn	Lys	Asp	Asn 265	Ile	Ala	Lys	Lys	Ala 270	Asn	Ser
	Ala	Asp	Va1 275	Tyr	Thr	Arg	Glu	Glu 280	Ser	Asp	Ser	Lys	Phe 285	٧a٦	Arg	Ile
10	Asp	G]y 290	Leu	Asn	Ala	Thr	Thr 295	Glu	Lys	Leu	Asp	Thr 300	Arg	Leu	Ala	ser
	А1а 305	Glu	Lys	Ser	Ile	А]а 310	Asp	Нis	Asp	Thr	Arg 315	Leu	Asn	Gly	Leu	Asp 320
15	Lys	Thr	val	Ser	Asp 325	Leu	Arg	Lys	Glu	Thr 330	Arg	Gln	Gly	Leu	Ala 335	Glu
	Gln	Ala	Ala	Leu 340	Ser	Gly	Leu	Phe	Gln 345	Pro	Tyr	Asn	Val	G]y 350	Gly	Ser
20	Gly	Gly	G]y 355	Gly	Thr	Ser	Ala	Pro 360	Asp	Phe	Asn	Аlа	Gly 365	Gly	Thr	Gly
	Ile	Gly 370	Ser	Asn	Ser	Arg	А1а 375	Thr	Thr	Αla	Lys	Ser 380	Ala	Аlа	val	Ser
25	Tyr 385	Ala	Gly	Ile	Lys	Asn 390	Glu	Met	Cys	Lys	Asp 395	Arg	Ser	Met	Leu	Cys 400
	Ala	Gly	Arg	Asp	Asp 405	Val	Ala	Val	Thr	Asp 410	Arg	Asp	Ala	Lys	Ile 415	Asn
30	Ala	Pro	Pro	Pro 420	Asn	Leu	His	Thr	Gly 425	Asp	Phe	Pro	Asn	Pro 430	Asn	Asp
	Аlа	Tyr	Lys 435	Asn	Leu	Ile	Asn	Leu 440	Lys	Pro	Ala	Ile	Glu 445	Ala	Gly	Tyr
35	Thr	Gly 450	Arg	Gly	va1	Glu	Val 455	Gly	Ile	val	Asp	Thr 460	Gly	Glu	Ser	٧a٦
	Gly 465	ser	Ile	Ser	Phe	Pro 470	Glu	Leu	Tyr	Gly	Arg 475	Lys	Glu	нis	Gly	туг 480
40	Asn	Glu	Asn	Tyr	Lys 485	Asn	Tyr	Thr	Ala	Tyr 490	Met	Arg	Lys	Glu	Ala 495	Pro
40	G∃u	Asp	Gly	G]y 500	Gly	Lys	Asp	Ile	G1u 505	Ala	Ser	Phe	Asp	Asp 510	Glu	Ala
45	Val	Ile	Glu 515	Thr	Glu	Ala	Lys	Pro 520	Thr	Asp	Ile	Arg	His 525	Val	Lys	Glu
45	Ile	G]y 530	His	Ile	Asp	Leu	va1 535	Ser	His	Ile	Ile	Gly 540	GТу	Arg	Ser	٧al
	Asp 545	Gly	Arg	Pro	Ala	G]y 550	Gly	ıle	Ala	Pro	Asp 555	Ala	Thr	Leu	His	11e 560
50	Met	Asn	Thr	Asn	Asp 565	Glu	Thr	Lys	Asn	G1u 570	Met	Met	val	Ala	Ala 575	Ile
	Arg	Asn	Ala	Trp 580	val	Lys	Leu	G⊺y	G]u 585	Arg	Gly	٧a٦	Arg	1]e 590	٧a٦	Asn
55	Asn	Ser	Phe 595	Gly	Thr	Thr	Ser	Arg 600	Ala	Gly	Thr	Ala	Asp 605	Leu	Phe	Gln

Ile Ala Asn Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser 610 615 620 Gly Gly Asp Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp 625 630 635 640 Tyr Gly Asn Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile 645 650 655 Phe Ser Thr Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu 660 665 670 10 Leu Pro Phe Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala 675 680 685 Gly Val Asp Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu 690 695 700 15 Pro Gly Thr Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr 705 710 715 720 Ala Met Trp Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr 725 730 735 20 Arg Thr Asn Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile 740 745 750Val Thr Gly Thr Ala Ala Leu Leu Leu Gln Lys Tyr Pro Trp Met Ser 755 760 765 25 Asn Asp Asn Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly 770 780 Ala Val Gly Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly 785 790 795 800 30 Lys Ala Met Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala 805 810 815 Asp Thr Lys Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile 820 825 83035 Ser Gly Thr Gly Gly Leu Ile Lys Lys Gly Gly Ser Gln Leu Gln Leu 835 840 845 His Gly Asn Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly Ser 850 855 860 40 Leu Val Leu Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys 865 870 880 Gly Ala Leu Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser 885 890 895 45 Asp Gly Ile Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu 900 905 910 Thr Val His Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu 915 920 925 50 Tyr Thr Arg Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile 930 940 Gly Gly Lys Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu 945 950 955 960 55 Asn Ser Thr Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly 965 970 975

	Gln	Asp	Tyr	ser 980	Phe	Phe	Thr	Asn	985	Glu	Thr	Asp	GIY	990	Leu	Leu
5	Ala	ser	Leu 995	Asp	Ser	٧a٦	Glu	Lys 1000		Ala	Gly	ser	Glu 1009		Asp	Thr
	Leu	Ser 1010		туг	٧a٦	Arg	Arg 101		Asn	Ala	Ala	Arg 1020		Аlа	Ser	Ala
10	Ala 1025		His	Ser	Ala	Pro 1030		Gly	Leu	Lys	Нis 1035		٧a٦	Glu	Gln	Gly 1040
	Gly	Ser	Asn	Leu	Glu 1045		Leu	Met	۷a٦	Glu 1050		Asp	Ala	Ser	Glu 1055	
15	Ser	Ala	Thr	Pro 1060		Thr	Val	Glu	Thr 1065		Ala	Ala	Asp	Arg 1070		Asp
	Met	Pro	Gly 1075	Ile 5	Arg	Pro	Tyr	Gly 1080		Thr	Phe	Arg	Ala 1085		Ala	Ala
20	val	Gln 1090		Ala	Asn	Ala	Ala 109		Gly	∨al	Arg	lle 1100		Asn	Ser	Leu
	Ala 110		Thr	۷al	Tyr	Ala 111(		Ser	Thr	Ala	Ala 1115		Ala	Asp	Met	Gln 1120
25	Gly	Arg	Arg	Leu	Lys 1125		val	Ser	Asp	Gly 1130		Asp	His	Asn	Gly 1135	
	Gly	Leu	Arg	Val 1140		Ala	Gln	Thr	Gln 1149		Asp	Gly	Gly	Thr 115(		Glu
30	Gln	Gly	Gly 115	Val	Glu	Gly	Lys	Met 1160		Gly	Ser	Thr	G]n 1165		۷al	Gly
	Ile	Ala 1170		Lys	Thr	Gly	Glu 1179		Thr	Thr	Ala	Ala 1180		Thr	Leu	Gly
35	Met 1185		Arg	Ser	Thr	⊤rp 1190		Glu	Asn	Ser	Ala 1199		Аlа	Lys	Thr	Asp 1200
	Ser	Ile	Ser	Leu	Phe 1205		Gly	Ile	Arg	His 121(		Ala	Gly	Asp	Ile 1215	Gly 5
40	Tyr	Leu	Lys	Gly 1220		Phe	Ser	Tyr	Gly 1225		Tyr	Lys	Asn	Ser 1230		Ser
	Arg	Ser	Thr 1235	Gly	Ala	Asp	Glu	His 1240		Glu	Gly	Ser	Val 1249	Asn	Gly	⊤hr
45	Leu	Met 1250		Leu	Gly	Ala	Leu 125		Gly	۷al	Asn	val 1260		Phe	Ala	Ala
	Thr 1269		Asp	Leu	Thr	Val 1270		Gly	Gly	Leu	Arg 127		Asp	Leu	Leu	Lys 1280
	Gln	Asp	Ala	Phe	Ala 1285		Lys	Gly	Ser	Ala 1290		Gly	Trp	Ser	Gly 129	
50	Ser	Leu	Thr	G]u 1300		Thr	Leu	٧a٦	Gly 1305		Ala	Gly	Leu	Lys 1310		Ser
	Gln	Pro	Leu 131	Ser	Asp	Lys	Ala	val 1320		Phe	Alа	Thr	Ala 1325		۷a٦	Glu
55	Arg	Asp 1330		Asn	Glу	Arg	Asp 133	туr	Thr	val	Thr	Gly 1340	Gly )	Phe	Thr	Gly

	Ala Thr 1345	Ala Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His Thr 1350 1355 1360	
5	Arg Leu	Val Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly Trp 1365 1370 1375	
	Asn Gly	Leu Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn 1380 1385 1390	
10	His Ser	Gly Arg Val Gly Val Gly Tyr Arg Phe 1395 1400	
15	<210> <211> <212> <213>	164 33 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cgcggato	164 ccg ctagcaaaac aaccgacaaa cgg	33
	<210> <211> <212> <213>	165 27 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cccgctcg	165 gag ttaccagcgg tagccta	27
30	<210> <211> <212> <213>	166 30 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> ctagctag	166 gcg gacacactta tttcggcatc	30
40	<210> <211> <212> <213>	167 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cccgctcg	167 gag ttaccagcgg tagcctaatt tg	32
50	<210> <211> <212> <213>	168 10 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> cccgctcg	168 gag	10
	<210>	169	

	<211> <212> <213>	33 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cgcggat	169 tccc atatgaaaac cttcttcaaa acc	33
10	<210> <211> <212> <213>	170 28 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cccgcto	170 cgag ttatttggct gcgccttc	28
20	<210> <211> <212> <213>	171 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> gcggcat	171 ttaa tatgttgaga aaattgttga aa <mark>tgg</mark>	35
30	<210> <211> <212> <213>	172 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> gcggcct	172 ccga gttattttt caaaatatat ttgc	34
	<210> <211> <212> <213>	173 33 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> gcggcca	173 atat gttacctaac cgtttcaaaa tgt	33
45	<210> <211> <212> <213>	174 31 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> gcggcct	174 ccga gttatttccg aggttttcgg g	31
55	<210> <211> <212> <213>	175 32 DNA Artificial Seguence	

```
<220>
                oligonucleotide
        <223>
        <400>
                                                                                  32
        cgcggatccc atatgacacg cttcaaatat tc
5
        <210>
                 176
        <211>
                 31
        <212>
                DNA
        <213>
                Artificial Sequence
10
        <220>
        <223>
                Oligonucleotide
        <400>
        cccgctcgag ttatttaaac cgataggtaa a
                                                                                   31
15
                 177
        <210>
        <211>
                 31
        <212>
                DNA
        <213>
                Artificial Sequence
        <220>
20
        <223>
                Oligonucleotide
        <400>
                 177
        cgcggatccc atatgggcag ggaagaaccg c
                                                                                   31
        <210>
                 178
25
        <211>
                 28
        <212>
<213>
                DNA
                Artificial Sequence
        <220>
<223>
                Oligonucleotide
30
        <400>
                 178
        gcccaagctt atcgatggaa tagccgcg
                                                                                   28
        <210>
                 179
        <211>
                 32
        <212>
35
                DNA
                Artificial Sequence
        <213>
        <220>
        <223>
                Oligonucleotide
        <400>
40
        cgcggatccg ctagcaacgg tttggatgcc cg
                                                                                   32
        <210>
                 180
        <211>
                 30
        <212>
                DNA
        <213>
                Artificial Sequence
45
        <220>
        <223>
                Oligonucleotide
        <400>
                180
                                                                                  30
        cccgctcgag tttgtctaag ttcctgatat
50
        <210>
                 181
        <211>
                 26
        <212>
                DNA
                Artificial Sequence
        <213>
55
        <220>
                Oligonucleotide
        <223>
```

	<400> cccgctc	181 gag attcccacct gccatc	26
5	<210> <211> <212> <213>	182 37 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cgcggat	182 ccg ctagcatgaa tttgcctatt caaaaat	37
15	<210> <211> <212> <213>	183 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cccgctc	183 gag ttaattccca cctgccatc	29
25	<210> <211> <212> <213>	184 37 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cgcggat	184 ccg ctagcatgaa tttgcctatt caaaaat	37
	<210> <211> <212> <213>	185 27 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cccgctc	185 gag ttggacgatg cccgcga	27
40	<210> <211> <212> <213>	186 37 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cgcggate	186 ccg ctagcatgaa tttgcctatt caaaaat	37
50	<210> <211> <212> <213>	187 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
<i>55</i>	<400>	187 gag ttattggacg atgcccgc	28

```
<210>
                 188
         <211>
                 32
         <212>
                 DNA
         <213>
                 Artificial Sequence
5
         <220>
         <223>
                 Oligonucleotide
         <400>
        cgcggatccc atatgtatcg caaactgatt gc
                                                                                   32
10
         <210>
                 189
         <211>
                 28
         <212>
                 DNA
         <213>
                 Artificial Sequence
         <220>
15
         <223>
                 Oligonucleotide
        <400>
                 189
        cccgctcgag ctaatcgatg gaatagcc
                                                                                   28
        <210>
                 190
20
        <211>
<212>
                 32
                 DNA
         <213>
                 Artificial Sequence
         <220>
        <223>
                 Oligonucleotide
25
        <400>
                                                                                   32
        cgcggatccc atatgaaaca gacagtcaaa tg
        <210>
                 191
        <211>
<212>
                 28
30
                 DNA
                 Artificial Sequence
        <213>
        <220>
                 Oligonucleotide
        <223>
35
        <400>
        cccgctcgag tcaataaccc gccttcag
                                                                                   28
        <210>
                 192
        <211>
                 45
         <212>
                 DNA
                 Artificial Sequence
40
        <213>
        <220>
        <223>
                 Oligonucleotide
        <400>
        cgcggatccc atatgttacg tttgactgct ttagccgtat gcacc
                                                                                   45
45
        <210>
                 193
        <211>
                 40
        <212>
                 DNA
        <213>
                 Artificial Sequence
50
        <220>
        <223>
                 Oligonucleotide
        <400>
                 193
        cccgctcgag ttattttgcc gcgttaaaag cgtcggcaac
                                                                                   40
55
        <210>
                 194
        <211>
                 32
```

	<212> <213>	DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cgcggate	194 ccc atatgaacaa aatataccgc at	32
10	<210> <211> <212> <213>	195 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cccgctc	195 gag ttaccactga taaccgac	28
20	<210> <211> <212> <213>	196 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cgcggate	196 ccc atatgaccga tgacgacgat ttat	34
	<210> <211> <212> <213>	197 28 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> gcccaage	197 ctt ccactgataa ccgacaga	28
35	<210> <211> <212> <213>	198 32 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400>	198 ccc atatgaacaa aatataccgc at	32
45	<210> <211> <212> <213>	199 28 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
50	<400> gcccaage	199 ctt ttaccactga taaccgac	28
55	<210> <211> <212> <213>	200 34 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> gggaatte	200 cca tatgggcatt tcccgcaaaa tatc	34
10	<210> <211> <212> <213>	201 40 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
15	<400> cccgctc	201 gag ttatttactc ctataacgag gtctcttaac	40
15	<210> <211> <212> <213>	202 36 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> gggaatte	202 cca tatgtcagat ttggcaaacg attctt	36
25	<210> <211> <212> <213>	203 40 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cccgctcg	203 gag ttatttactc ctataacgag gtctcttaac	40
35	<210> <211> <212> <213>	204 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> gggaatte	204 cca tatgggcatt tcccgcaaaa tatc	34
45	<210> <211> <212> <213>	205 32 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
50		205 gag ttacgtatca tatttcacgt gc	32
	<210> <211> <212> <213>	206 35 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> gggaat	206 tcca tatgcacgtg aaatatgata cgaag	35
5	<210> <211> <212> <213>	207 37 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
10	<400> cccgct	207 cgag tttactccta taacgaggtc tcttaac	37
15	<210> <211> <212> <213>	208 36 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> gggaat	208 tcca tatgtcagat ttggcaaacg attctt	36
25	<210> <211> <212> <213>	209 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cccgct	209 cgag cgtatcatat ttcacgtgc	29
30	<210> <211> <212> <213>	210 36 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> gggaat	210 tcca tatgtcagat ttggcaaacg attctt	36
40	<210> <211> <212> <213>	211 37 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
,,,	<400> cccgct	211 cgag tttactccta taacgaggtc tcttaac	37
50	<210> <211> <212> <213>	212 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> cgcgga	212 tccc atatgcaaaa tgcgttcaaa atccc	35

5	<210> <211> <212> <213>	213 32 DNA Artificial Sequence	
3	<220> <223>	Oligonucleotide	
10		213 ccc atatgaacaa aatataccgc at	32
		214 29 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cccgctcg	214 gag tttgctttcg atagaacgg	29
20		215 34 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
20	<400> gcggccat	215 at ggtcataaaa tatacaaatt tgaa	34
30	<210> <211> <212> <213>	216 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> gcggcctc	216 :ga gttagcctga gacctttgca aatt	34
40	<212>	217 30 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
45	<400> gcggccat	217 at gaaacagaaa aaaaccgctg	30
45		218 32 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> gcggcctc	218 ga gttacggttt gacaccgttt tc	32
55	<210> <211> <212>	219 30 DNA	

	<213>	Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cgcggate	219 ccc atatgaaaac cctgctcctc	30
10	<210> <211> <212> <213>	220 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cccgctcg	220 gag ttatcctcct ttgcggc	27
20	<210> <211> <212> <213>	221 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> gcggccat	221 tat ggcaaaaatg atgaaatggg	30
	<210> <211> <212> <213>	222 29 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> gcggccto	222 cga gttatcggcg cggcgggcc	29
35	<210> <211> <212> <213>	223 30 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> gcggccat	223 tat gaaaaaatcc tccctcatca	30
45	<210> <211> <212> <213>	224 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> gcggcctd	224 cga gttatttgcc gccgtttttg gc	32
	<210> <211> <212> <213>	225 31 DNA Artificial Sequence	
55	<220>	·	

	<223>	Oligonucleotide	
5	<400> gcggcca	225 tat ggcccctgcc gacgcggtaa g	31
	<210> <211> <212> <213>	226 33 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> gcggccte	226 cga gtttgccgcc gtttttggct ttc	33
15	<210> <211> <212> <213>	227 30 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> gcggcca	227 tat gaaacacata ctccccctga	30
25	<210> <211> <212> <213>	228 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> gcggccte	228 cga gttattcgcc tacggttttt tg	32
35	<210> <211> <212> <213>	229 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> gcggcca	229 tat gatttacatc gtactgtttc	30
	<210> <211> <212> <213>	230 32 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> gcggccto	230 ga gttaggagaa caggcgcaat gc	32
50	<210> <211> <212> <213>	231 32 DNA Artificial Sequence	
<i>55</i>	<220> <223>	Oligonucleotide	
	<400>	231	

	gcggcca	tat gtacaacatg tatcaggaaa ac	32
5	<210> <211> <212> <213>	232 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> gcggcct	232 cga gggagaacag gcgcaatgcg g	31
15	<210> <211> <212> <213>	233 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cgcggate	233 ccg ctagctgcgg cacggcggg	29
	<210> <211> <212> <213>	234 28 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cccgctcg	234 gag ataacggtat gccgccag	28
30	<210> <211> <212> <213>	235 31 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cgcggate	235 ccc atatggaatc aacactttca c	31
40	<210> <211> <212> <213>	236 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cccgctcg	236 gag ttacacgcgg ttgctgt	27
50	<210> <211> <212> <213>	237 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
<i>55</i>	<400> cgcggato	237 ccc atatgaacaa cagacatttt g	31
	<210>	238	

	<211> <212> <213>	28 DNA Artificial Sequence	
5	<220> <223>	oligonucleotide	
	<400> cccgctc	238 cgag ttacctgtcc ggtaaaag	28
10	<210> <211> <212> <213>	239 33 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cgcggat	239 ccg ctagcaccgt catcaaacag gaa	33
20	<210> <211> <212> <213>	240 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cccgctc	240 gag tcaagattcg acgggga	27
30	<210> <211> <212> <213>	241 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cgcggat	241 ccc atatgtccgc aaacgaatac g	31
	<210> <211> <212> <213>	242 28 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cccgctc	242 gag tcagtgttct gccagttt	28
45	<210> <211> <212> <213>	243 29 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cgcggat	243 ccc atatgccgtc tgaaacacg	29
55	<210> <211> <212> <213>	244 28 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> cccgct	244 cgag ttagcggagc agtttttc	28
10	<210> <211> <212> <213>	245 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cgcgga	245 tccc atatgaccgc catcagcc	28
70	<210> <211> <212> <213>	246 27 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgct	246 cgag ttaaagccgg gtaacgc	27
25	<210> <211> <212> <213>	247 31 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> gcggcca	247 atat ggaaacacag ctttacatcg g	31
35	<210> <211> <212> <213>	248 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> gcggcct	248 tcga gtcaataata atatcccgcg	30
45	<210> <211> <212> <213>	249 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> gcggcca	249 atat gattaaaatc cgcaatatcc	30
	<210> <211> <212> <213>	250 36 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> gcggcct	250 cga gttaaatctt ggtagattgg atttgg	36
5	<210> <211> <212> <213>	251 30 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
10	<400> gcggcca	251 tat gactgacaac gcactgctcc	30
15	<210> <211> <212> <213>	252 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> gcggcct	252 cga gtcagaccgc gttgtcgaaa c	31
25	<210> <211> <212> <213>	253 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cgcggat	253 ccc atatggcgtt aaaaacatca aa	32
	<210> <211> <212> <213>	254 27 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cccgctc	254 gag tcagcccttc atacagc	27
40	<210> <211> <212> <213>	255 32 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> gcggcat	255 taa tggcacaaac tacactcaaa cc	32
50	<210> <211> <212> <213>	256 33 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> gcggcct	256 cga gttaaaactt cacgttcacg ccg	33

-	<210> <211> <212> <213>	257 34 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> gcggcat	257 taa tgcatgaaac tgagcaatcg gtgg	34
10	<210> <211> <212> <213>	258 38 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> gcggcct	258 cga gaaacttcac gttcacgccg ccggtaaa	38
20	<210> <211> <212> <213>	259 33 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cgcggate	259 ccc atatgggcaa atccgaaaat acg	33
30	<210> <211> <212> <213>	260 26 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cccgctc	260 gag ataatggcgg cggcgg	26
40	<210> <211> <212> <213>	261 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cgcggate	261 ccc atatgtttcc ccccgacaa	29
	<210> <211> <212> <213>	262 31 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cccgctc	262 gag tcattctgta aaaaaagtat g	31
55	<210> <211>	263 32	

	<212> <213>	DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cgcggat	263 ccc atatgcttca aagcgacagc ag	32
10	<210> <211> <212> <213>	264 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cccgctc	264 gag ttcggatttt tgcgtactc	29
20	<210> <211> <212> <213>	265 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cgcggat	265 ccc atatggcaat ggcagaaaac g	31
30	<210> <211> <212> <213>	266 27 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cccgctc	266 gag ctatacaatc cgtgccg	27
35	<210> <211> <212> <213>	267 32 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cgcggat	267 ccc atatggattc tttttcaaa cc	32
45	<210> <211> <212> <213>	268 27 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cccgctc	268 gag tcagttcaga aagcggg	27
55	<210> <211> <212> <213>	269 32 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> cgcggat	269 ccc atatgaaacc tttgatttta gg	32
	<210> <211> <212> <213>	270 28 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cccgctc	270 gag ttatttgggc tgctcttc	28
15	<210> <211> <212> <213>	271 30 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cgcggat	271 ccc atatggtaat cgtctggttg	30
25	<210> <211> <212> <213>	272 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cccgctc	272 gag ctacgacttg gttaccg	27
35	<210> <211> <212> <213>	273 33 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> gcggcca	273 tat gagacgtaaa atgctaaagc tac	33
45	<210> <211> <212> <213>	274 31 DNA Artificial Sequence	
70	<220> <223>	Oligonucleotide	
50		274 cga gtcaaagtgt tctgtttgcg c	31
	<210> <211> <212> <213>	275 30 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> gccgcca	275 Itat gttgacttta acccgaaaaa	30
5	<210> <211> <212> <213>	276 34 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
10	<400> gccgcct	276 cga ggccggcggt caataccgcc cgaa	34
15	<210> <211> <212> <213>	277 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cgcggat	277 ccc atatggcgca atgcgatttg ac	32
25	<210> <211> <212> <213>	278 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
00	<400> cccgctc	278 gag.ttcggcggta aatgccg	27
30	<210> <211> <212> <213>	279 28 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> gcggcca	279 tat ggcggggccg atttttgt	28
40	<210> <211> <212> <213>	280 33 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
43	<400> gcggcct	280 cga gttatttgct ttcagtatta ttg	33
50	<210> <211> <212> <213>	281 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> gcggcca	281 tat gaactttgct ttatccgtca	30

5	<210> <211> <212> <213>	282 32 DNA Artificial Sequence	
Ŭ	<220> <223>	Oligonucleotide	
10	<400> gcggccto	282 cga gttaacggca gtatttgttt ac	32
70	<210> <211> <212> <213>	283 31 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cgcggato	283 ccc atatgggttt gcgcttcggg c	31
20	<210> <211> <212> <213>	284 29 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
20	<400> gcccaago	284 ctt ttttcctttg ccgtttccg	29
30	<210> <211> <212> <213>	285 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cgcggato	285 ccc atatggccga cctttccgaa aa	32
40	<210> <211> <212> <213>	286 27 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cccgctcg	286 gag gaagcgcgtt cccaagc	27
45	<210> <211> <212> <213>	287 29 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cgcggato	287 ccc atatgcacga cacccgtac	29
55	<210> <211> <212>	288 28 DNA	

```
<213>
                  Artificial Sequence
          <220>
                  Oligonucleotide
          <223>
5
          <400>
                                                                                    28
          cccgctcgag ttagaagcgc gttcccaa
          <210>
                  289
          <211>
                  34
10
          <212>
                  DNA
          <213>
                  Artificial Sequence
          <220>
          <223>
                  Oligonucleotide
          <400>
15
          ctagctagct ttaaacgcag cgtaatcgca atgg
                                                                                    34
          <210>
          <211>
<212>
                  31
                  DNA
                  Artificial Sequence
          <213>
20
          <220>
          <223>
                  Oligonucleotideucleotide
          <400>
                  290
          cccgctcgag tcaatcctgc tcttttttgc c
                                                                                    31
25
          <210>
                  25
          <211>
          <212>
                  DNA
          <213>
                  Artificial Sequence
          <220>
30
          <223>
                  Oligonucleotide
          <400>
          ctagctagcg ggggcggcgg tggcg
                                                                                    25
          <210>
                  292
35
                  31
          <211>
          <212>
                  Artificial Sequence
          <213>
          <220>
                  Oligonucleotide
          <223>
40
          <400>
                                                                                    31
          cccgctcgag tcaatcctgc tcttttttgc c
          <210>
                  293
          <211>
                  40
          <212>
45
                  DNA
          <213>
                  Artificial Sequence
          <220>
          <223>
                  oligonucleotide
50
                                                                                    40
          ctagctagcg ctcatcctcg ccgcctgcgg gggcggcggt
          <210>
                  294
                  31
          <211>
          <212>
          <213>
                  Artificial Sequence
55
          <220>
```

	<223>	Oligonucleotide	
5	<400> cccgctc	294 gag tcaatcctgc tcttttttgc c	31
	<210> <211> <212> <213>	295 25 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cggggate	295 ccg ggggcggcgg tggcg	25
15	<210> <211> <212> <213>	296 31 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgctcg	296 gag tcaatcctgc tcttttttgc c	31
25	<210> <211> <212> <213>	297 25 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> ctagctag	297 gcg ggggcggcgg tggcg	25
35	<210> <211> <212> <213>	298 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cccgctcg	298 gag atcctgctct tttttgcc	28
	<210> <211> <212> <213>	299 28 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> ctagctag	299 gct gcgggggcgg cggtggcg	28
50	<210> <211> <212> <213>	300 28 DNA Artificial Sequence	
<i>55</i>	<220> <223>	Oligonucleotide	
	<400>	300	

	cccgctcg	gag atcctgctct tttttgcc	28
5	<210> <211> <212> <213>	301 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> cgcggato	301 ccg ctagccccga tgttaaatcg gc	32
15	<210> <211> <212> <213>	302 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cgcggato	302 ccg ctagccaaga tatggcggca gt	32
	<210> <211> <212> <213>	303 32 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cgcggato	303 ccg ctagcgccga atccgcaaat ca	32
30	<210> <211> <212> <213>	304 32 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cgcgctag	304 gcg gaagggttga tttggctaat gg	32
40	<210> <211> <212> <213>	305 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cgcgctag	305 gcg gaagggttga tttggctaat gg	32
50	<210> <211> <212> <213>	306 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> cgccatat	306 tgt ttaaacgcag cgtaatcgc	29
	<210>	307	

	<211> <212> <213>	34 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cccgctcg	307 gag aaaattgcta ccgccattcg cagg	34
10	<210> <211> <212> <213>	308 32 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cgccata	308 tgg gaagggttga tttggctaat gg	32
20	<210> <211> <212> <213>	309 38 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cccgctcg	309 gag cttgtcttta taaatgatga catatttg	38
30	<210> <211> <212> <213>	310 40 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cccgctc	310 gag tttataaaag ataatatatt gattgattcc	40
	<210> <211> <212> <213>	311 31 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cgcgctag	311 gca tgccgctgat tcccgtcaat c	31
45	<210> <211> <212> <213>	312 25 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> ctagctag	312 gcg ggggcggcgg tggcg	25
<i>55</i>	<210> <211> <212> <213>	313 31 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> cccgctc	313 gag tcaatcctgc tcttttttgc c	31
10	<210> <211> <212> <213>	314 32 DNA Artificial Sequence	
70	<220> <223>	Oligonucleotide	
15	<400> cgcggat	314 ccg ctagccccga tgttaaatcg gc	32
	<210> <211> <212> <213>	315 28 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgctc	315 gag atcctgctct tttttgcc	28
25	<210> <211> <212> <213>	316 32 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cgcggat	316 ccg ctagccccga tgttaaatcg gc	32
35	<210> <211> <212> <213>	317 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cccgctc	317 gag tcaatcctgc tcttttttgc c	31
45	<210> <211> <212> <213>	318 87 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> cgcggate gcctgtt	318 ccg ctagctttga acgcagtgtg attgcaatgg cttgtatttt tgccctttca cgc ccgatgttaa atcggcg	60 87
55	<210> <211> <212> <213>	319 31 DNA Artificial Sequence	
	<220>		

	<223>	Oligonucleotide	
5	<400> cccgctc	319 gag tcaatcctgc tctttttgc c	31
	<210> <211> <212> <213>	320 90 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
		320 ccg ctagcaaaac cttcttcaaa accctttccg ccgccgcact cgcgctcatc cct gctcgcccga tgttaaatcg	60 90
15	<210> <211> <212> <213>	321 31 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgctc	321 gag tcaatcctgc tctttttgc c	31
25	<210> <211> <212> <213>	322 33 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
30	<400> cgcggat	322 ccc atatgaaaac caagttaatc aaa	33
35	<210> <211> <212> <213>	323 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cccgctc	323 gag ttattgattt ttgcggatga	30
45	<210> <211> <212> <213>	324 34 DNA Artificial Sequence	
43	<220> <223>	Oligonucleotide	
50	<400> cgcggat	324 ccc atatgttaaa tcgggtattt tatc	34
	<210> <211> <212> <213>	325 28 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> cccgctc	325 gag ttaatccgcc attccctg	28
5	<210> <211> <212> <213>	326 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> gcggcca	326 Itat gaaattacaa caattggctg	30
15	<210> <211> <212> <213>	327 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> gcggcct	327 cga gttaccttac gtttttcaaa g	31
25	<210> <211> <212> <213>	328 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cgcggat	328 ccc atatgcaagc acggctgct	29
50	<210> <211> <212> <213>	329 29 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cccgcto	329 gag tcaaggttgt ccttgtcta	29
40	<210> <211> <212> <213>	330 30 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cgcggat	330 ccc atatgatgaa accgcacaac	30
50	<210> <211> <212> <213>	331 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> cccgcto	331 gag tcagttgctc aacacgtc	28

5	<210> <211> <212> <213>	332 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> cgcggate	332 ccc atatggtaga cgcgcttaag ca	32
	<210> <211> <212> <213>	333 25 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cccgctc	333 gag agctgcatgg cggcg	25
20	<210> <211> <212> <213>	334 30 DNA Artificial Sequence	
<i>25</i>	<220> <223>	Oligonucleotide	
	<400> cgcggate	334 ccc atatggcacg gtcgttatac	30
30	<210> <211> <212> <213>	335 26 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cccgctc	335 gag ctaccgcgca ttcctg	26
	<210> <211> <212> <213>	336 31 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> gcggccat	336 tat ggaatttttc attatcttgt t	31
45	<210> <211> <212> <213>	337 31 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> gcggccto	337 cga gttatttggc ggttttgctg c	31
55	<210> <211> <212>	338 32 DNA	

	<213>	Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> gcggccat	338 tat gaagtatgtc cggttatttt tc	32
10	<210> <211> <212> <213>	339 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> gcggccto	339 cga gttatcggct tgtgcaacgg	30
20	<210> <211> <212> <213>	340 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cgcggato	340 ccg ctagctccgg cagcaaaacc ga	32
	<210> <211> <212> <213>	341 28 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> gcccaago	341 ctt acgcagttcg gaatggag	28
35	<210> <211> <212> <213>	342 35 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> gccgccat	342 cat gttgaatatt aaactgaaaa ccttg	35
45	<210> <211> <212> <213>	343 32 DNA . Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> gccgccto	343 cga gttatttctg atgccttttc cc	32
<i>55</i>	<210> <211> <212> <213>	344 29 DNA Artificial Sequence	
	<220>		

	<223>	Oligonucleotide	
5	<400> gccgcca	344 tat ggacaataag accaaactg	29
9	<210> <211> <212> <213>	345 30 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> gccgcct	345 cga gttaacggtg cggacgtttc	30
15	<210> <211> <212> <213>	346 32 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cgcggat	346 ccc atatgaacaa actgtttctt ac	32
25	<210> <211> <212> <213>	347 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cccgctc	347 gag tcattccgcc ttcagaaa	28
35	<210> <211> <212> <213>	348 45 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cgcggate	348 ccc atatgcaagg tatcgttgcc gacaaatccg cacct	45
	<210> <211> <212> <213>	349 42 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cccgctc	349 gag agctaattgt gcttggtttg cagataggag tt	42
50	<210> <211> <212> <213>	350 52 DNA Artificial Sequence	
<i>55</i>	<220> <223>	Oligonucleotide	
	<400>	350	

```
cgcggatccc atatgaaccg caccctgtac aaagttgtat ttaacaaaca tc
                                                                                   52
         <210>
         <211>
<212>
                  45
5
                 DNA
         <213>
                 Artificial Sequence
         <220>
<223>
                 Oligonucleotide
         <400>
10
         cccgctcgag ttaagctaat tgtgcttggt ttgcagatag gagtt
                                                                                   45
         <210>
                  352
                  46
         <211>
         <212>
                 DNA
                 Artificial Sequence
         <213>
15
         <220>
         <223>
                 Oligonucleotide
         <400>
         cgcggatccc atatgacggg agaaaatcat gcggtttcac ttcatg
                                                                                   46
20
         <210>
         <211>
<212>
                  42
                 DNA
                 Artificial Sequence
         <213>
25
         <220>
                 Oligonucleotide
         <223>
         <400>
         cccgctcgag agctaattgt gcttggtttg cagataggag tt
                                                                                   42
         <210>
30
         <211>
                  51
         <212>
                 DNA
                 Artificial Sequence
         <213>
         <220>
         <223>
                 Oligonucleotide
35
         <400>
         cgcggatccc atatggtttc agacggccta tacaaccaac atggtgaaat t
                                                                                   51
         <210>
                  355
         <211>
                  41
40
         <212>
         <213>
                 Artificial Sequence
         <220>
         <223>
                 Oligonucleotide
         <400>
45
         cccgctcgag gcggtaactg ccgcttgcac tgaatccgta a
                                                                                   41
         <210>
                 356
         <211>
                 46
         <212>
                 DNA
         <213>
                 Artificial Sequence
50
         <220>
                 Oligonucleotide
         <223>
         <400>
         cgcggatccc atatgacggg agaaaatcat gcggtttcac ttcatg
                                                                                   46
55
         <210>
                 357
```

	<211> <212> <213>	41 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cccgct	357 cgag gcggtaactg ccgcttgcac tgaatccgta a	41
10	<210>	358	
	<211> <212>	49 DNA	
	<213>	Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cgcgga	358 tccc atatgcaaag caaagtcaaa gcagaccatg cctccgtaa	49
	<210>	359	
20	<211> <212>	56 DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
25	<400> cccgct	359 cgag tcttttcctt tcaattataa ctttagtagg ttcaattttg gtcccc	56
	<210>	360	
	<211> <212>	51 DNA	
30	<213>	Artificial Sequence	
	<220> <223>	Oligonucleotide	
	<400>	360	
35	cgcgga	tccc atatggtttc agacggccta tacaaccaac atggtgaaat t	51
	<210>	361	
	<211> <212>	56 DNA	
	<213>	Artificial Sequence	
40	<220>		
	<223>	Oligonucleotide	
	<400>	361	5.6
		cgag tcttttcctt tcaattataa ctttagtagg ttcaattttg gtcccc	56
45	<210> <211>	362 27	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
50	<223>	Oligonucleotide	
	<400>	362 atat gacccgtttg acccgcg	27
			۷,
	<210> <211>	363 31	
55	<212>	DNA	
	<213>	Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> gcggcct	363 cga gtcagcgggc gttcatttct t	31
10	<210> <211> <212> <213>	364 33 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cgcggat	364 ccc atatgaacac cattttcaaa atc	33
	<210> <211> <212> <213>	365 32 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgctc	365 gag ttaatttact tttttgatgt cg	32
25	<210> <211> <212> <213>	366 28 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> gcggcca	366 tat ggattcgccc aaggtcgg	28
35	<210> <211> <212> <213>	367 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> gcggcct	367 cga gctacacttc ccccgaagtg g	31
45	<210> <211> <212> <213>	368 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> cgcggat	368 ccc atatgatagt tgaccaaagc c	31
	<210> <211> <212> <213>	369 30 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> cccgct	369 cgag ttattttcc gatttttcgg	30
5	<210> <211> <212> <213>	370 28 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
10	<400> gcggcc	370 atat gcttgaactg aacggact	28
15	<210> <211> <212> <213>	371 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> gcggcc	371 tcga gtcagcggaa gcggacgatt	30
25	<210> <211> <212> <213>	372 34 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
30	<400> cgcgga	372 tccc atatgtccaa actcaaaacc atcg	34
	<210> <211> <212> <213>	373 29 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cccgct	373 cgag gcttccaatc agtttgacc	29
40	<210> <211> <212> <213>	374 32 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> gcggcc	374 atat gagcgcaatc gttgatattt tc	32
50	<210> <211> <212> <213>	375 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400>	375 tcga gttatttgcc cagttggtag aatg	34

5	<210> <211> <212> <213>	376 32 DNA Artificial Sequence	
Ü	<220> <223>	Oligonucleotide	
10	<400> gcggcca	376 atat ggtgatacat ccgcactact tc	32
	<210> <211> <212> <213>	377 32 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> gcggcct	377 tcga gtcaaaatcg agttttacac ca	32
20	<210> <211> <212> <213>	378 31 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> gcggcca	378 atat gaccatctat ttcaaaaacg g	31
30	<210> <211> <212> <213>	379 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> gcggcct	379 tcga gtcagccgat gtttagcgtc catt	34
40	<210> <211> <212> <213>	380 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cgcggat	380 tccc atatgagcag cggaggggt g	31
	<210> <211> <212> <213>	381 27 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cccgcto	381 cgag ttgcttggcg gcaaggc	27
55	<210> <211>	382 31	

	<212> <213>	DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cgcggat	382 ccc atatggtcgc cgccgacatc g	31
10	<210> <211> <212> <213>	383 27 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
7.5	<400> cccgctc	383 gag ttgcttggcg gcaaggc	27
20	<210> <211> <212> <213>	384 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cgcggat	384 ccc atatgggcgg ttcggaaggc g	31
30	<210> <211> <212> <213>	385 33 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cccgctc	385 gag ttgaacactg atgtctttc cga	33
55	<210> <211> <212> <213>	386 35 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cgcggat	386 ccg ctagcaaact gtcgttggtg ttaac	35
45	<210> <211> <212> <213>	387 26 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cccgctc	387 gag ttgacccgct ccacgg	26
55	<210> <211> <212> <213>	388 31 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> gccgcca	388 tat ggcggacttg gcgcaagacc c	31
10	<210> <211> <212> <213>	389 39 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
15	<400> gccgccto	389 cga gatctcctaa acctgtttta acaatgccg	39
73	<210> <211> <212> <213>	390 31 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> gccgcca1	390 tat ggcggacttg gcgcaagacc c	31
25	<210> <211> <212> <213>	391 31 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> gcggccto	391 cga gctccatgct gttgccccag c	31
35	<210> <211> <212> <213>	392 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> gccgcca1	392 tat ggcggacttg gcgcaagacc c	31
45	<210> <211> <212> <213>	393 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> gcggcctd	393 cga gaaaatcccc gctaaccgca g	31
	<210> <211> <212> <213>	394 31 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> cgcggat	394 ccc atatgagcag cggagggggt g	31
5	<210> <211> <212> <213>	395 27 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
10	<400> cccgctc	395 gag ttgcttggcg gcaaggc	27
15	<210> <211> <212> <213>	396 31 DNA Artificial Sequence	
cgcgga  5		Oligonucleotide	
20		396 ccc atatggtcgc cgccgacatc g	31
25	<211> <212>	397 27 DNA Artificial Sequence	
		Oligonucleotide	
		397 gag ttgcttggcg gcaaggc	27
210> 211> 212> 213>  220> 223>   400> 2211> 212> 211> 212> 213>  220> 2213>  220> 2213>  220> 2213>  2210> 2213> 2213>  220> 2213> 2213> 220> 2213> 2213> 220> 2213> 2213> 220> 2213> 2213> 220> 2213> 220> 2213> 2213> 220> 2213> 2213> 220> 2213> 220> 2213> 2213> 220> 2213> 2213> 220> 2213> 2213> 220> 2213> 2213> 220> 2213> 2213> 220> 2213> 2213> 220> 2213>	398 33 DNA Artificial Sequence		
35		Oligonucleotide	
		398 ccc atatggacgg tgttgtgcct gtt	33
40	<211> <212>	399 29 DNA Artificial Sequence	
		Oligonucleotide	
45		399 gag cttacggatc aaattgacg	29
50	<211> <212>	400 33 DNA Artificial Sequence	
		Oligonucleotide	
55		400 ccc atatgggcag ccaatctgaa gaa	33

E	<210> <211> <212> <213>	401 28 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
10	<400> cccgct	401 cgag ctcagctttt gccgtcaa	28
70	<210> <211> <212> <213>	402 33 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cgcgga	402 tccg ctagctactc atccattgtc cgc	33
20	<210> <211> <212> <213>	403 29 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
23	<400> cccgct	403 cgag ccagttgtag cctattttg	29
30	<210> <211> <212> <213>	404 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cgcgga	404 tccg ctagcatgcg cttcacacac ac	32
	<210> <211> <212> <213>	405 30 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
30 35 40 45	<400> cccgcte	405 cgag ttaccagttg tagcctattt	30
45	<210> <211> <212> <213>	406 32 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> gccgcca	406 atat ggcacaaacg gaaggtttgg aa	32
55	<210> <211> <212>	407 36 DNA	

```
<213>
                 Artificial Sequence
         <220>
         <223>
                 Oligonucleotide
5
         <400>
                                                                                   36
         gccgcctcga gaaaactgta acgcaggttt gccgtc
         <210>
                  408
         <211>
                 32
10
         <212>
         <213>
                 Artificial Sequence
         <220>
<223>
                 Oligonucleotide
15
         <400>
                 408
         gcggccatat ggaagaaaca ccgcgcgaac cg
                                                                                   32
                 409
         <211>
<212>
                 32
                 DNA
         <213>
                 Artificial Sequence
20
         <220>
         <223>
                 Oligonucleotide
         <400>
        gcggcctcga ggaacgtttt attaaactcg ac
                                                                                   32
25
                 410
         <210>
         <211>
         <212>
                 DNA
                 Artificial Sequence
         <213>
         <220>
<223>
30
                 Oligonucleotide
         <400>
                                                                                   32
        gcggccatat gagaaaaccg accgataccc ta
                 411
33
         <210>
35
         <211>
         <212>
                 Artificial Sequence
         <213>
         <220>
         <223>
                 oligonucleotide
40
        <400>
                 411
         gcggcctcga gtcaacgcca ctgccagcgg ttg
                                                                                   33
         <210>
                 412
         <211>
                 48
         <212>
45
                 DNA
                 Artificial Sequence
         <213>
         <220>
                 Oligonucleotide
         <223>
50
         cgcggatccc atatgaagaa gaacatattg gaattttggg tcggactg
                                                                                   48
         <210>
                 413
         <211>
                 38
         <212>
                 DNA
         <213>
                 Artificial Sequence
55
         <220>
```

	<400>		
5		413 gag ttattcggcg gctttttccg cattgccg	38
5	<210> <211> <212> <213>	414 103 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
		414 cca tatgaaaaag acagctatcg cgattgcagt ggcactggct ggtttcgcta gca ggccgctagc gctttccgcg tggccggcgg tgc	60 103
15	<210> <211> <212> <213>	415 38 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgctc	415 gag ttattcggcg gctttttccg cattgccg	38
25	<210> <211> <212> <213>	416 32 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> catgcca	416 tgg ctttccgcgt ggccggct gc	32
35	<210> <211> <212> <213>	417 38 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cccgctc	417 gag ttattcggcg gctttttccg cattgccg	38
45	<210> <211> <212> <213>	418 31 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
50		418 ccc atatgtttgc cgaaacccgc c	31
	<210> <211> <212> <213>	419 28 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> cccgctc	419 gag aggttgtgtt ccaggttg	28
5	<210> <211> <212> <213>	420 31 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cgcggat	420 ccc atatgaaaaa aaccgcctat g	31
15	<210> <211> <212> <213>	421 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cccgctc	421 gag ttaaggttgt gttccagg	28
25	<210> <211> <212> <213>	422 33 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cgcggat	422 ccc atatgaaaaa atacctattc cgc	33
	<210> <211> <212> <213>	423 27 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cccgctc	423 gag ttacgggcgg tattcgg	27
40	<210> <211> <212> <213>	424 34 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cgcggat	424 ccc atatgcaaag caagagcatc caaa	34
50	<210> <211> <212> <213>	425 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> cccgctc	425 gag ttacgggcgg tattcgg	27

-	<210> <211> <212> <213>	426 86 DNA Artificial Sequence	
5	<220> <223>	oligonucleotide	
10		426 cca tatgaaaacc ttcttcaaaa ccctttccgc cgccgcgcta gcgctcatcc ctg ccaaagcaag agcatc	60 86
	<210> <211> <212> <213>	427 38 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cccgctcg	427 gag ttacgggcgg tattcgggct tcataccg	38
20	<210> <211> <212> <213>	428 33 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cgcggato	428 ccg tcgactgtgg gggcggcggt ggc	33
30	<210> <211> <212> <213>	429 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cccgctcg	429 gag tcaatcctgc tctttttgc c	31
40	<210> <211> <212> <213>	430 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> gcggccat	430 tat gaagaaaaca ttgacactgc	30
	<210> <211> <212> <213>	431 32 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> gcggccto	431 cga gttaatggtg cgaatgaccg at	32
55	<210> <211>	432 46	

	<212> <213>	DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> ggggaca	432 aagt ttgtacaaaa aagcaggctt gcggcaagga tgccgg	46
10	<210> <211> <212> <213>	433 47 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> ggggaco	433 cact ttgtacaaga aagctgggtc taaagcaaca atgccgg	47
20	<210> <211> <212> <213>	434 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cgcggat	434 cccc atatgaaaca caccgtatcc	30
	<210> <211> <212> <213>	435 26 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cccgcto	435 cgag ttatctcgtg cgcgcc	26
35	<210> <211> <212> <213>	436 30 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cgcggat	436 ccc atatgagccc cgcgccgatt	30
45	<210> <211> <212> <213>	437 28 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
50	<400> cccgcto	437 gag tttttgtgcg gtcaggcg	28
<i>55</i>	<210> <211> <212> <213>	438 62 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> ggggacaa tc	438 agt ttgtacaaaa aagcaggctt gttcgtttgg gggatttaaa ccaaaccaaa	60 62
10	<210> <211> <212> <213>	439 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cgcggate	439 ccc atatggcgga tgcgcccgcg	30
	<210> <211> <212> <213>	440 26 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgctcg	440 gag aaaccgccaa tccgcc	26
25	<210> <211> <212> <213>	441 61 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> ggggacca t	441 act ttgtacaaga aagctgggtt cattttgttt ttccttcttc tcgaggccat	60 61
35	<210> <211> <212> <213>	442 30 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cgcggato	442 ccc atatgaaacc caaaccgcac	30
45	<210> <211> <212> <213>	443 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> cccgctcg	443 gag tcagcgttgg acgtagt	27
<i>55</i>	<210> <211> <212> <213>	444 33 DNA Artificial Sequence	
	<220>		

	<223>	Oligonucleotide	
5	<400> gggaatte	444 cca tatgaaaaaa atcatcttcg ccg	33
9	<210> <211> <212> <213>	445 31 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cccgctcg	445 gag ttattgtttg gctgcctcga t	31
15	<210> <211> <212> <213>	446 33 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> gggaatte	446 cca tatggccacc tacaaagtgg acg	33
25	<210> <211> <212> <213>	447 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cggggate	447 cct tgtttggctg cctcgatttg	30
35	<210> <211> <212> <213>	448 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cgcggate	448 ccc atatgcaaga acaatcgcag aaag	34
	<210> <211> <212> <213>	449 30 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cccgctcg	449 gag ttttttcggc aaattggctt	30
50	<210> <211> <212> <213>	450 45 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	
	<400>	450	

	ggggaca	agt ttgtacaaaa aagcaggctg ccgatgccgt tgcgg	45
5	<210> <211> <212> <213>	451 47 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> ggggacca	451 act ttgtacaaga aagctgggtt cagggtcgtt tgttgcg	47
15	<210> <211> <212> <213>	452 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cgcggate	452 ccc atatgaaaca ctttccatcc	30
	<210> <211> <212> <213>	453 28 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cccgctcg	453 gag ttaccactcg taattgac	28
30	<210> <211> <212> <213>	454 30 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cgcggate	454 ccc atatggccac aagcgacgac	30
40	<210> <211> <212> <213>	455 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cccgctcg	455 gag ttaccactcg taattgac	28
50	<210> <211> <212> <213>	456 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
<i>55</i>	<400> cgcggato	456 ccc atatggccac aaacgacg	28
	<210>	457	

	<211> <212> <213>	28 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cccgcto	457 cgag acccacgttg taaggttg	28
10	<210> <211> <212> <213>	458 32 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cgcgga	458 tccc atatggccac aagcgacgac ga	32
20	<210> <211> <212> <213>	459 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cccgcto	459 cgag acccacgttg taaggttg	28
30	<210> <211> <212> <213>	460 33 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
<i>35</i>	<400> cgcgga1	460 tccc atatgatgaa acactttcca tcc	33
	<210> <211> <212> <213>	461 29 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cccgcto	461 cgag ttaacccacg ttgtaaggt	29
45	<210> <211> <212> <213>	462 33 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cgcgga1	462 tccc atatgatgaa acactttcca tcc	33
55	<210> <211> <212> <213>	463 29 DNA Artificial Seguence	

	<220> <223>	Oligonucleotide	
5	<400> cccgcte	463 cgag ttaacccacg ttgtaaggt	29
	<210> <211>	464 28	
10	<212> <213>	DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cgcgga	464 tccc atatggccac aaacgacg	28
	<210> <211>	465 29	
	<212> <213>	DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgcto	465 cgag gtctgacact gttttatcc	29
25	<210> <211>	466 33	
	<212> <213>	DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cgcgga1	466 tccc atatgatgaa acactttcca tcc	33
<i>35</i>	<210> <211>	467 29	
33	<212> <213>	DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cccgcto	467 cgag ttatgctttg gcggcaaag	29
	<210> <211>	468 30	
45	<212> <213>	DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> cgcgga1	468 tccc atatggccac aaacgacgac	30
	<210> <211>	469 27	
	<212> <213>	DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> cgcgga	469 tccc cactcgtaat tgacgcc	27
5	<210> <211> <212> <213>	470 30 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cgcgga	470 tccc atatggccac aagcgacgac	30
15	<210> <211> <212> <213>	471 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cgcgga	471 tccc cactcgtaat tgacgcc	27
25	<210> <211> <212> <213>	472 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cgcgga	472 tccc atatggccac aaacgacgac	30
	<210> <211> <212> <213>	473 27 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cgcgga	473 tcca cccacgttgt aaggttg	27
40	<210> <211> <212> <213>	474 33 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cgcgga	474 tccc atatgatgaa acactttcca tcc	33
50	<210> <211> <212> <213>	475 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> cgcgga	475 tcca cccacgttgt aaggttg	27

	<210> <211> <212> <213>	476 25 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
10	<400> cgcgga	476 tccg gagggggtgg tgtcg	25
10	<210> <211> <212> <213>	477 27 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cccgcte	477 cgag ttgcttggcg gcaaggc	27
20	<210> <211> <212> <213>	478 25 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cgcgga	478 tccg gcggaggcgg cactt	25
30	<210> <211> <212> <213>	479 26 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cccgcte	479 cgag gaaccggtag cctacg	26
40	<210> <211> <212> <213>	480 41 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cgcgga	480 tccg gtggtggtgg ttcagatttg gcaaacgatt c	41
	<210> <211> <212> <213>	481 29 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cccgcto	481 cgag cgtatcatat ttcacgtgc	29
55	<210> <211>	482 25	

	<212> <213>	DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cgcggat	482 ccg gagggggtgg tgtcg	25
10	<210> <211> <212> <213>	483 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cccgctc	483 gag ttattgcttg gcggcaag	28
20	<210> <211> <212> <213>	484 25 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cgcggat	484 ccg gcggaggcgg cactt	25
	<210> <211> <212> <213>	485 28 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cccgctc	485 gag tcagaaccgg tagcctac	28
35	<210> <211> <212> <213>	486 41 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cgcggat	486 ccg gtggtggtgg ttcagatttg gcaaacgatt c	41
45	<210> <211> <212> <213>	487 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> cccgctc	- 487 gag ttacgtatca tatttcacgt gc	32
<i>55</i>	<210> <211> <212> <213>	488 31 DNA Artificial Sequence	

```
<220>
<223>
                  Oligonucleotide
         <400>
5
         cgcggatccc atatggccac aagcgacgac g
                                                                                      31
         <210>
                  489
         <211>
         <212>
                  DNA
         <213>
                  Artificial Sequence
10
         <220>
<223>
                  Oligonucleotide
         <400>
         cccgctcgag ccactcgtaa ttgacgcc
                                                                                      28
15
         <210>
                  490
         <211>
<212>
                  30
                  DNA
         <213>
                  Artificial Sequence
         <220>
<223>
20
                  Oligonucleotide
         <400>
                  490
         cgcggatccc atatggccac aaacgacgac
                                                                                      30
         <210>
25
         <211>
<212>
                  28
                  DNA
         <213>
                  Artificial Sequence
         <220>
         <223>
                  Oligonucleotide
30
         <400>
         cccgctcgag tgctttggcg gcaaagtt
                                                                                      28
         <210>
                  492
         <211>
                  30
         <212>
                  DNA
35
                  Artificial Sequence
         <213>
         <220>
         <223>
                  Oligonucleotide
         <400>
40
         cgcggatccc atatggccac aaacgacgac
                                                                                      30
         <210>
         <211>
<212>
                  37
                 DNA
         <213>
                 Artificial Sequence
45
         <220>
         <223>
                 Oligonucleotide
         <400>
         cccgctcgag tttagcaata ttatctttgt tcgtagc
                                                                                      37
50
         <210>
                  494
         <211>
                  32
         <212>
                 DNA
         <213>
                 Artificial Sequence
         <220>
55
         <223>
                 Oligonucleotide
```

	<400> cgcggate	494 ccc atatgaaagc aaaccgtgcc ga	32
5	<210> <211> <212> <213>	495 28 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
10	<400> cccgctcg	495 gag ccactcgtaa ttgacgcc	28
15	<210> <211> <212> <213>	496 61 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> ggggacaa c	496 agt ttgtacaaaa aagcaggctg cagccacaaa cgacgacgat gttaaaaaaag	60 61
25	<210> <211> <212> <213>	497 61 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> ggggacca g	497 act ttgtacaaga aagctgggtt taccactcgt aattgacgcc gacatggtag	60 61
<i>35</i>	<210> <211> <212> <213>	498 31 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> gcggcca1	498 tat ggcagcaaaa gacgtacagt t	31
	<210> <211> <212> <213>	499 33 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> gcggccto	499 cga gttacatcat gccgcccata cca	33
50	<210> <211> <212> <213>	500 31 DNA Artificial Sequence	
<i>55</i>	<220> <223>	Oligonucleotide	
	<400>	500	

	cgcggat	ccg ctagcttagg cggcggcgga g	31
5	<210> <211> <212> <213>	501 26 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> cccgctcg	501 gag gaaccggtag cctacg	26
15	<210> <211> <212> <213>	502 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cccctage	502 cta gcacttctgc gcccgactt	29
	<210> <211> <212> <213>	503 26 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cccgctc	503 gag gaaccggtag cctacg	26
30	<210> <211> <212> <213>	504 31 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cgcggate	504 ccg ctagcttagg cggcggcgga g	31
40	<210> <211> <212> <213>	505 26 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cccgctc	505 gag gaaccggtag cctacg	26
50	<210> <211> <212> <213>	506 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> cgcggate	506 ccg ctagcacttc tgcgcccgac tt	32
	<210>	507	

	<211> <212> <213>	26 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cccgct	507 cgag gaaccggtag cctacg	26
10	<210> <211> <212> <213>	508 50 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cgcgga	508 tccg ctagccgaac gaccccaacc ttccctacaa aaactttcaa	50
20	<210> <211> <212> <213>	509 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cccgct	509 cgag tcagaaccga cgtgccaagc cgttc	35
30	<210> <211> <212> <213>	510 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> gccgcca	510 atat gcccccactg gaagaacgga cg	32
	<210> <211> <212> <213>	511 35 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> gccgcct	511 tcga gtaataaacc ttctatgggc agcag	35
45	<210> <211> <212> <213>	512 31 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cgcgga1	512 tccc atatgtccgt ccacgcatcc g	31
55	<210> <211> <212> <213>	513 31 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> cccgcto	513 gag tttgaatttg taggtgtatt g	31
	<210> <211> <212> <213>	514 29 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
15	<400> cgcggat	514 ccc atatgacccc ttccgcact	29
15	<210> <211> <212> <213>	515 32 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgcto	515 gag ttatttgaat ttgtaggtgt at	32
25	<210> <211> <212> <213>	516 33 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cgcggat	516 ccc atatgaaaac caattcagaa gaa	33
35	<210> <211> <212> <213>	517 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cccgctc	517 gag tccacagaga ttgtttcc	28
45	<210> <211> <212> <213>	518 17 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50		518 gaa gggcggg	17
	<210> <211> <212> <213>	519 29 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> gcccaa	519 gctt tcagaagaag acttcacgc	29
5	<210> <211> <212> <213>	520 36 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cgcgga	520 tccc atatgcaaac ccataaatac gctatt	36
15	<210> <211> <212> <213>	521 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> gcccaa	521 gctt gaagaagact tcacgccag	29
25	<210> <211> <212> <213>	522 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cgcgga	522 tccc atatggtctt tttcgacaat accga	35
	<210> <211> <212> <213>	523 10 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> gcccaa	523 gctt	10
40	<210> <211> <212> <213>	524 36 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cgcgga	524 tccc atatgaataa aactttaaaa aggcgg	36
50	<210> <211> <212> <213>	525 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> gcccaa	525 gctt tcagaagaag acttcacgc	29

5	<210> <211> <212> <213>	526 35 DNA Artificial Sequence	
J	<220> <223>	Oligonucleotide	
10	<400> cgcgaate	526 ccc atatgttcga tcttgattct gtcga	35
	<210> <211> <212> <213>	527 28 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cccgctcg	527 gag tcgcacaggc tgttggcg	28
20	<210> <211> <212> <213>	528 32 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cgcgaato	528 ccc atatgttggg cggaggcggc ag	32
30	<210> <211> <212> <213>	529 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cccgctcg	529 gag tcgcacaggc tgttggcg	28
40	<210> <211> <212> <213>	530 32 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cgcgaate	530 ccc atatgttggg cggaggcggc ag	32
	<210> <211> <212> <213>	531 28 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cccgctcg	531 gag tcgcacaggc tgttggcg	28
55	<210> <211>	532 33	

	<212> <213>	DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cgcgga1	532 tccc atatggcaaa tttggaggtg cgc	33
10	<210> <211> <212> <213>	533 27 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cccgcto	533 cgag ttcggagcgg ttgaagc	27
20	<210> <211> <212> <213>	534 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cgcgga1	534 tccc atatgcaacg tcgtattata accc	34
	<210> <211> <212> <213>	535 29 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cccgcto	535 cgag ttattcggag cggttgaag	29
35	<210> <211> <212> <213>	536 42 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cgcggat	536 tccc atatgggcat caaagtcgcc atcaacggct ac	42
45	<210> <211> <212> <213>	537 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> cccgcto	537 cgag tttgagcggg cgcacttcaa gtccg	35
55	<210> <211> <212> <213>	538 33 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> cgcggat	538 cccc atatgggcgg cagcgaaaaa aac	33
	<210> <211> <212> <213>	539 28 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cccgcto	539 cgag gttggtgccg actttgat	28
15	<210> <211> <212> <213>	540 31 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cgcggat	540 cccc atatgggcgg cggaagcgat a	31
25	<210> <211> <212> <213>	541 27 DNA Artificial Sequence	
00	<220> <223>	Oligonucleotide	
30	<400> cccgcto	541 cgag tttgcccgct ttgagcc	27
35	<210> <211> <212> <213>	542 33 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cgcggat	542 ccc atatgggcaa atccgaaaat acg	33
	<210> <211> <212> <213>	543 27 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
50	<400> cccgcto	543 gag catcccgtac tgtttcg	27
	<210> <211> <212> <213>	544 62 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> ggggacaa aa	544 agt ttgtacaaaa aagcaggctc cgacattacc gtgtacaacg gccaacaaag	60 62
5	<210> <211> <212> <213>	545 61 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> ggggacca g	545 act ttgtacaaga aagctgggtc ttatttcata ccggcttgct caagcagccg	60 61
15	<210> <211> <212> <213>	546 61 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> ggggacaa a	546 agt ttgtacaaaa aagcaggctg atacggtgtt ttcctgtaaa acggacaaca	60 61
25	<210> <211> <212> <213>	547 60 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
30	<400> ggggacca	547 act ttgtacaaga aagctgggtc taggaaaaat cgtcatcgtt gaaattcgcc	60
35	<210> <211> <212> <213>	548 47 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> ggggaca	548 agt ttgtacaaaa aagcaggcta tgcaccccat cgaaacc	47
45	<210> <211> <212> <213>	549 47 DNA Artificial Sequence	
43	<220> <223>	Oligonucleotide	
50	<400> ggggacca	549 act ttgtacaaga aagctgggtc tagtcttgca gtgcctc	47
	<210> <211> <212> <213>	550 41 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	

	<400> cgcggat	550 ccc atatgggaaa tttcttatat agaggcatta g	41	
5	<210> <211> <212> <213>	551 40 DNA Artificial Sequence		
10	<220> <223>	Oligonucleotide		
10	<400> cccgctc	551 gag gttaatttct atcaactctt tagcaataat	40	
15	<210> <211> <212> <213>	552 31 DNA Artificial Sequence		
	<220> <223>	Oligonucleotide		
20	<400> cgcggat	552 ccc atatggcctg ccaagacgac a	31	
25	<210> <211> <212> <213>	553 26 DNA Artificial Sequence		
	<220> <223>	Oligonucleotide		
30	<400> 553 cccgctcgag ccgcctcctg ccgaaa			
	<210> <211> <212> <213>	554 34 DNA Artificial Sequence		
35	<220> <223>	Oligonucleotide		
	<400> cgcggat	554 ccc atatggcaga gatctgtttg ataa	34	
40	<210> <211> <212> <213>	555 27 DNA Artificial Sequence		
45	<220> <223>	Oligonucleotide		
	<400> cccgctc	555 gag cggttttccg cccaatg	27	
50	<210> <211> <212> <213>	556 30 DNA Artificial Sequence		
	<220> <223>	Oligonucleotide		
55	<400> cgcggat	556 ccc atatgcagcc ggatacggtc	30	

5	<210> <211> <212> <213>	557 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> cccgcto	557 cgag aatcacttcc aacacaaaat	30
	<210> <211> <212> <213>	558 33 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cgcgga1	558 tccc atatgtggtt gctgatgaag ggc	33
20	<210> <211> <212> <213>	559 28 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cccgcto	559 cgag gactgcttca tcttctgc	28
30	<210> <211> <212> <213>	560 34 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cgcggat	560 tccc atatggaact gatgactgtt ttgc	34
40	<210> <211> <212> <213>	561 29 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
45	<400> cccgcto	561 cgag tcagactgct tcatcttct	29
45	<210> <211> <212> <213>	562 45 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> cgcggat	562 tccc atatgagcat taaagtagcg attaacggtt tcggc	45
55	<210> <211> <212>	563 40 DNA	

```
<213>
                 Artificial Sequence
         <220>
                 Oligonucleotide
         <223>
5
        <400>
        cccgctcgag gattttgcct gcgaagtatt ccaaagtgcg
                                                                                   40
         <210>
                 564
         <211>
                 32
         <212>
10
                 DNA
         <213>
                 Artificial Sequence
         <220>
         <223>
                 Oligonucleotide
         <400>
15
        cgcggatccg ctagccccga tgttaaatcg gc
                                                                                   32
                 565
29
         <210>
         <211>
         <212>
                 DNA
                 Artificial Sequence
         <213>
20
         <220>
        <223>
                 Oligonucleotide
         <400>
        cggggatcca tcctgctctt ttttgccgg
                                                                                   29
25
         <210>
                 566
                 42
         <211>
         <212>
                 DNA
         <213>
                 Artificial Sequence
         <220>
30
                 Oligonucleotide
         <223>
         <400>
                                                                                   42
        cgcggatccg gtggtggtgg tcaaagcaag agcatccaaa cc
         <210>
                 567
35
        <211>
                 30
        <212>
                 DNA
                 Artificial Sequence
        <213>
         <220>
         <223>
                 Oligonucleotide
40
        <400>
                                                                                   30
        cccaagcttt tcgggcggta ttcgggcttc
                 568
39
        <210>
         <211>
         <212>
                 DNA
45
        <213>
                 Artificial Sequence
        <220>
        <223>
                 Oligonucleotide
        <400>
50
        cgcggatccg gtggtggtgg tgccacctac aaagtggac
                                                                                   39
        <210>
                 569
                 28
        <211>
         <212>
         <213>
                 Artificial Sequence
55
        <220>
```

	<223>	Oligonucleotide	
5	<400> gcccaage	569 ctt ttgtttggct gcctcgat	28
J	<210> <211> <212> <213>	570 34 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cgcggate	570 ccg gtggtggtgg tacaagcgac gacg	34
15	<210> <211> <212> <213>	571 28 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> gcccaage	571 ctt ccactcgtaa ttgacgcc	28
25	<210> <211> <212> <213>	572 41 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> cgcggate	572 ccg gtggtggtgg ttcagatttg gcaaacgatt c	41
35	<210> <211> <212> <213>	573 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cccaagc	573 ttc gtatcatatt tcacgtgc	28
	<210> <211> <212> <213>	574 44 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cccaagct	574 ttg gtggtggtgg tggttcagat ttggcaaacg attc	44
50	<210> <211> <212> <213>	575 29 DNA Artificial Sequence	
55	<220> <223>	Oligonucleotide	
55	<400>	575	

	cccgctcg	gag cgtatcatat ttcacgtgc	29
5	<210> <211> <212> <213>	576 45 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
10	<400> cccaagct	576 ttg gtggtggtgg tggtcaaagc aagagcatcc aaacc	45
15	<210> <211> <212> <213>	577 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cccgctcg	577 gag cgggcggtat tcgggctt	28
	<210> <211> <212> <213>	578 32 DNA Artificial Sequence	
25	<220> <223>	Oligonucleotide	
	<400> cgcggato	578 ccg ctagccccga tgttaaatcg gc	32
30	<210> <211> <212> <213>	579 29 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> cggggato	579 cca tcctgctctt ttttgccgg	29
40	<210> <211> <212> <213>	580 36 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> cgcggato	580 ccg ctagcggaca cacttatttc ggcatc	36
50	<210> <211> <212> <213>	581 30 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
<i>55</i>	<400> cgcggato	581 cc cagcggtagc ctaatttgat	30
55	<210>	582	

	<211> <212> <213>	41 DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cgcggate	582 ccg gtggtggtgg ttcagatttg gcaaacgatt c	41
10	<210> <211> <212> <213>	583 28 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cccaagct	583 ttc gtatcatatt tcacgtgc	28
20	<210> <211> <212> <213>	584 36 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> gcggcgto	584 cga cggtggcgga ggcactggat cctcag	36
30	<210> <211> <212> <213>	585 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> ggaggca	585 ctg gatcctcaga tttggcaaac gattc	35
	<210> <211> <212> <213>	586 29 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> cccgctcg	586 gag cgtatcatat ttcacgtgc	29
45	<210> <211> <212> <213>	587 33 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> ggaattco	587 cat atgtcagatt tggcaaacga ttc	33
55	<210> <211> <212> <213>	588 28 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> cgcggat	588 ccc gtatcatatt tcacgtgc	28
10	<210> <211> <212> <213>	589 25 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> cggggat	589 ccg ggggcggcgg tggcg	25
	<210> <211> <212> <213>	590 30 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccaago	590 tta tcctgctctt ttttgccggc	30
25	<210> <211> <212> <213>	591 42 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> cgcggat	591 ccg gtggtggtgg tcaaagcaag agcatccaaa cc	42
35	<210> <211> <212> <213>	592 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> cccaago	592 ttc gggcggtatt cgggcttc	28
45	<210> <211> <212> <213>	593 26 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> ccccaag	593 ctt gggggcggcg gtggcg	26
	<210> <211> <212> <213>	594 31 DNA Artificial Sequence	
55	<220>	Oligonucleotide	

	<400> cccgcto	594 cgag atcctgctct tttttgccgg c	31
5	<210> <211> <212> <213>	595 45 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> cccaago	595 cttg gtggtggtgg tggtcaaagc aagagcatcc aaacc	45
15	<210> <211> <212> <213>	596 28 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
20	<400> cccgcto	596 gag cgggcggtat tcgggctt	28
25	<210> <211> <212> <213>	597 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
30	<400> ggaggca	597 actg gatccgcagc cacaaacgac gacga	35
	<210> <211> <212> <213>	598 36 DNA Artificial Sequence	
35	<220> <223>	Oligonucleotide	
	<400> gcggcct	598 ccga gggtggcgga ggcactggat ccgcag	36
40	<210> <211> <212> <213>	599 28 DNA Artificial Sequence	
45	<220> <223>	Oligonucleotide	
	<400> cccgcto	599 gag acccagcttg taaggttg	28
50	<210> <211> <212> <213>	600 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
55	<400> ggaggca	600 actg gatccgcagc cacaaacgac gacga	35

5	<210> <211> <212> <213>	601 36 DNA Artificial Sequence	
3	<220> <223>	Oligonucleotide	
10	<400> gcggcct	601 cga gggtggcgga ggcactggat ccgcag	36
10	<210> <211> <212> <213>	602 28 DNA Artificial Sequence	
15	<220> <223>	Oligonucleotide	
	<400> cccgctc	602 gag ccactcgtaa ttgacgcc	28
20	<210> <211> <212> <213>	603 38 DNA Artificial Sequence	
<i>25</i>	<220> <223>	Oligonucleotide	
	<400> gcggcct	603 cga gggatccggc ggaggcggca cttctgcg	38
30	<210> <211> <212> <213>	604 26 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
35	<400> cccgctc	604 gag gaaccggtag cctacg	26
40	<210> <211> <212> <213>	605 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
45	<400> ggaggca	605 actg gatcctcaga tttggcaaac gattc	35
	<210> <211> <212> <213>	606 37 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
	<400> gcggcgt	- 606 cga cggtggcgga ggcactggat cctcaga	37
55	<210> <211>	607 29	

	<212> <213>	DNA Artificial Sequence	
5	<220> <223>	Oligonucleotide	
	<400> cccgctc	607 gag cgtatcatat ttcacgtgc	29
10	<210> <211> <212> <213>	608 35 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
15	<400> gcggccte	608 cga gggatccgga gggggtggtg tcgcc	35
20	<210> <211> <212> <213>	609 25 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
25	<400> cccgctc	609 gag ttgcttggcg gcaag	25
20	<210> <211> <212> <213>	610 35 DNA Artificial Sequence	
30	<220> <223>	Oligonucleotide	
	<400> ggaggca	610 ctg gatccgcagc cacaaacgac gacga	35
35	<210> <211> <212> <213>	611 36 DNA Artificial Sequence	
40	<220> <223>	Oligonucleotide	
	<400> gcggccte	611 cga gggtggcgga ggcactggat ccgcag	36
45	<210> <211> <212> <213>	612 28 DNA Artificial Sequence	
50	<220> <223>	Oligonucleotide	
55	<400> cccgctcg	612 gag acccagcttg taaggttg	28
<i>55</i>	<210> <211> <212> <213>	613 35 DNA Artificial Sequence	

	<220> <223>	Oligonucleotide	
5	<400> ggaggca	613 ctg gatccgcagc cacaaacgac gacga	35
	<210> <211> <212> <213>	614 36 DNA Artificial Sequence	
10	<220> <223>	Oligonucleotide	
	<400> gcggcct	614 cga gggtggcgga ggcactggat ccgcag	36
15	<210> <211> <212> <213>	615 28 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
	<400> cccgctc	615 gag ccactcgtaa ttgacgcc	28
25	<210> <211> <212> <213>	616 35 DNA Artificial Sequence	
20	<220> <223>	Oligonucleotide	
30	<400> ggaggca	616 ctg gatcctcaga tttggcaaac gattc	35
35	<210> <211> <212> <213>	617 37 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
40	<400> gcggcgt	617 cga cggtggcgga ggcactggat cctcaga	37
45	<210> <211> <212> <213>	618 29 DNA Artificial Sequence	
	<220> <223>	Oligonucleotide	
50	<400> cccgctc	618 gag cgtatcatat ttcacgtgc	29
	<210> <211> <212> <213>	619 488 PRT Artificial Sequence	
55	<220> <223>	MC58	

Met Phe Lys Arg Ser Val Ile Ala Met Ala Cys Ile Phe Ala Leu Ser 1 10 15 Ala Cys Gly Gly Gly Gly Gly Ser Pro Asp Val Lys Ser Ala Asp 20 25 30 Thr Leu Ser Lys Pro Ala Ala Pro Val Val Ser Glu Lys Glu Thr Glu 35 40 45 Ala Lys Glu Asp Ala Pro Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro 50 55 60 10 Ser Ala Gln Gly Ser Gln Asp Met Ala Ala Val Ser Glu Glu Asn Thr 65 75 80 Gly Asn Gly Gly Ala Val Thr Ala Asp Asn Pro Lys Asn Glu Asp Glu 85 90 95 15 Val Ala Gln Asn Asp Met Pro Gln Asn Ala Ala Gly Thr Asp Ser Ser 100 105 110Thr Pro Asn His Thr Pro Asp Pro Asn Met Leu Ala Gly Asn Met Glu 115 120 125 20 Asn Gln Ala Thr Asp Ala Gly Glu Ser Ser Gln Pro Ala Asn Gln Pro 130 140 Asp Met Ala Asn Ala Asp Gly Met Gln Gly Asp Asp Pro Ser Ala 145 150 155 25 Gly Gly Gln Asn Ala Gly Asn Thr Ala Ala Gln Gly Ala Asn Gln Ala 165 170 Gly Asn Asn Gln Ala Ala Gly Ser Ser Asp Pro Ile Pro Ala Ser Asn 180 185 190 30 Pro Ala Pro Ala Asn Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala 195 200 205 Asn Gly Val Leu Ile Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His 210 220 35 Cys Lys Gly Asp Ser Cys Ser Gly Asn Asn Phe Leu Asp Glu Glu Val 235 235 Gln Leu Lys Ser Glu Phe Glu Lys Leu Ser Asp Ala Asp Lys Ile Ser 245 250 255 40 Asn Tyr Lys Lys Asp Gly Lys Asn Asp Lys Phe Val Gly Leu Val Ala 260 270 Asp Ser Val Gln Met Lys Gly Ile Asn Gln Tyr Ile Ile Phe Tyr Lys 275 280 285 45 Pro Lys Pro Thr Ser Phe Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg 290 295 300 Arg Ser Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp 305 310 315 32050 Thr Leu Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly 325 330 335 Asn Ile Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala  $340 \hspace{1cm} 345 \hspace{1cm} 350$ 55 Glu Lys Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro 355 360 365

	Ala Lys 370	Gly Glu	Met		Ala 375	Gly	Ala	ΑΊа	۷a٦	Tyr 380	Asn	Gly	Glu	val
5	Leu His 385	Phe His		Glu / 390	Asn	Gly	Arg	Pro	Tyr 395	Pro	Thr	Arg	Glу	Arg 400
	Phe Ala	Ala Lys	val . 405	Asp	Phe	Gly	Ser	Lys 410	Ser	val	Asp	Gly	11e 415	ıle
10	Asp Ser	Gly Asp 420	Asp	Leu	His	Met	G]y 425	Thr	Gln	Lys	Phe	Lys 430	Ala	Ala
	Ile Asp	Gly Asn 435	Gly	Phe	Lys	Gly 440	Thr	Тгр	Thr	Glu	Asn 445	Gly	Ser	Gly
15	Asp Val 450	Ser Gly	Lys		Tyr 455	Gly	Pro	Аlа	Gly	G]u 460	Glu	۷al	Ala	Gly
	Lys Tyr 465	Ser Tyr		Pro 470	Thr	Asp	Ala	Glu	Lys 475	Glу	Glу	Phe	Gly	val 480
20	Phe Ala	Gly Lys	Lys 485	Glu (	G∏n	Asp								
	<210> <211> <212> <213>	620 427 PRT Artific	ial S	eque	nce									
25	<220> <223>	2996												
30	<400> Met Phe 1	620 Glu Arg	Ser 5	Val :	ıle	Ala	Met	А]а 10	Cys	Ile	Phe	Ala	Leu 15	Ser
	Ala Cys	Gly Gly 20	Gly	Gly (	Gly	Gly	Ser 25	Pro	Asp	۷a٦	Lys	Ser 30	Ala	Asp
35	Thr Leu	Ser Lys 35	Pro ,	Ala /	Ala	Pro 40	Val	٧a٦	Ala	Glu	Lys 45	Glu	Thr	Glu
	Val Lys 50	Glu Asp	Ala		Gln 55	Ala	Gly	Ser	Gln	Gly 60	G∏n	Gly	Ala	Pro
40	Ser Thr 65	Gln Gly		Gln / 70	Asp	Met	Ala	Ala	va1 75	Ser	Ala	G]u	Asn	Thr 80
	Gly Asn	Gly Gly	Ala . 85	Ala '	Thr	Thr	Asp	Lys 90	Pro	Lys	Asn	Glu	Asp 95	Glu
45	Gly Pro	Gln Asn 100		Met	Pro	Gln	Asn 105	Ser	Ala ·	Glu	Ser	Ala 110	Asn	Gln
	Thr Gly	Asn Asn 115	Gln	Pro /	Ala	Asp 120	Ser	Ser	Asp	Ser	Ala 125	Pro	ΑΊα	Ser
50	Asn Pro 130	Ala Pro	Ala .	Asn (	Gly 135	Gly	Ser	Asn	Phe	Gly 140	Arg	Val	Asp	Leu
	Ala Asn 145	Gly Val		11e / 150	Asp	Gly	Pro	Ser	G]n 155	Asn	Ile	Thr	Leu	Thr 160
55	His Cys	Lys Gly	Asp 165	ser (	Cys	Asn	Gly	Asp 170	Asn	Leu	Leu	Asp	Glu 175	Glu
55	Ala Pro	Ser Lys 180		Glu	Phe	Glu	Asn 185	Leu	Asn	Glu	Ser	Glu 190	Arg	Ile

Glu Lys Tyr Lys Lys Asp Gly Lys Ser Asp Lys Phe Thr Asn Leu Val 195 200 205 Ala Thr Ala Val Gln Ala Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr 210 215 220 Lys Asp Lys Ser Ala Ser Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala 225 230 235 Arg Ser Arg Ser Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn 245 250 Leu 25510 Gln Ala Asp Thr Leu Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly 260 265 270 His Ser Gly Asn Ile Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr 275 280 285 Tyr Gly Ala Glu Lys Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln 290 295 300 Gly Glu Pro Ala Lys Gly Glu Met Leu Ala Gly Thr Ala Val Tyr Asn 305 310 31520 Gly Glu Val Leu His Phe His Thr Glu Asn Gly Arg Pro Tyr Pro Thr 325 330 335Arg Gly Arg Phe Ala Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp 340 345 35025 Gly Ile Ile Asp Ser Gly Asp Asp Leu His Met Gly Thr Gln Lys Phe  $355 \\ \hspace*{1.5cm} 360 \\ \hspace*{1.5cm} 365$ Lys Ala Ala Ile Asp Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn 370 375 380 30 Gly Gly Gly Asp Val Ser Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu 385 390 395 400 Val Ala Gly Lys Tyr Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly 405 410 415 35 Phe Gly Val Phe Ala Gly Lys Lys Glu Gln Asp 420 425

#### Claims

40

45

50

- 1. A method for the heterologous expression of a protein of the invention, in which (a) at least one domain in the protein is deleted and, optionally, (b) no fusion partner is used.
- 2. The method of claim 1, in which the protein of the invention is ORF46.
- 3. The method of claim 2, in which ORF46 is divided into a first domain (amino acids 1-433) and a second domain (amino acids 433-608).
- 4. The method of claim 2, in which the protein of the invention is 564.
- 5. The method of claim 4, in which protein 564 is divided into domains as shown in Figure 8.
- 55 **6.** The method of claim 1 in which the protein of the invention is 961.
  - 7. The method of claim 6, in which protein 961 is divided into domains as shown in Figure 12.

- 8. The method of claim 1, in which the protein of the invention is 502 and the domain is amino acids 28 to 167 (numbered according to the MC58 sequence).
- 9. The method of claim 1, in which the protein of the invention is 287.

10

15

20

30

35

40

45

50

- 10. A method for the heterologous expression of a protein of the invention, in which (a) a portion of the N-terminal domain of the protein is deleted.
- 11. The method of claim 9 or claim 10, in which protein 287 is divided into domains A B & C shown in Figure 5.
- 12. The method of claim 11, in which (i) domain A, (ii) domains A and B, or (iii) domains A and C are deleted.
- 13. The method of claim 11, wherein (i) amino acids 1-17, (ii) amino acids 1-25, (iii) amino acids 1-69, or (iv) amino acids 1-106, of domain A are deleted.
- **14.** A method for the heterologous expression of a protein of the invention, in which (a) no fusion partner is used, and (b) the protein's native leader peptide (if present) is used.
- **15.** The method of claim 14, in which the protein of the invention is selected from the group consisting of: 111, 149, 206, 225-1, 235, 247-1, 274, 283, 286, 292, 401, 406, 502-1, 503, 519-1, 525-1, 552, 556, 557, 570, 576-1, 580, 583, 664, 759, 907, 913, 920-1, 936-1, 953, 961, 983, 989, Orf4, Orf7-1, Orf9-1, Orf23, Orf25, Orf37, Orf38, Orf40, Orf40.1, Orf40.2, Orf72-1, Orf76-1, Orf85-2, Orf91, Orf97-1, Orf119, Orf143.1, NMB0109, NMB2050, 008, 105, 117-1, 121-1, 122-1, 128-1, 148, 216, 243, 308, 593, 652, 726, 926, 982, Orf83-1 and Orf143-1.
- 25 16. A method for the heterologous expression of a protein of the invention, in which (a) the protein's leader peptide is replaced by the leader peptide from a different protein and, optionally, (b) no fusion partner is used.
  - 17. The method of claim 16, in which the different protein is 961, ORF4, *E.coli* OmpA, or *E.carotovora* PelB, or in which the leader peptide is MKKYLFSAA.
  - 18. The method of claim 17, in which the different protein is E.coli OmpA and the protein of the invention is ORF1.
  - **19.** The method of claim 17, in which the protein of the invention is 911 and the different protein is *E.carotovora* PelB or *E.coli* OmpA.
  - 20. The method of claim 17, in which the different protein is ORF4 and the protein of the invention is 287.
  - 21. A method for the heterologous expression of a protein of the invention, in which (a) the protein's leader peptide is deleted and, optionally, (b) no fusion partner is used.
  - 22. The method of claim 21, in which the protein of the invention is 919.
  - 23. A method for the heterologous expression of a protein of the invention, in which expression of a protein of the invention is carried out at a temperature at which a toxic activity of the protein is not manifested.
  - **24.** The method of claim 23, in which protein 919 is expressed at 30°C.
  - **25.** A method for the heterologous expression of a protein of the invention, in which protein is mutated to reduce or eliminate toxic activity.
  - 26. The method of claim 25, in which the protein of the invention is 907, 919 or 922.
  - 27. The method of claim 26, in which 907 is mutated at Glu-117 (e.g. Glu→Gly).
- 55 **28.** The method of claim 26, in which 919 is mutated at Glu-255 (e.g. Glu $\rightarrow$ Gly) and/or Glu-323 (e.g. Glu $\rightarrow$ Gly).
  - **29.** The method of claim 26, in which 922 is mutated at Glu-164 (*e.g.* Glu→Gly), Ser-213 (*e.g.* Ser→Gly) and/or Asn-348 (*e.g.* Asn→Gly).

- **30.** A method for the heterologous expression of a protein of the invention, in which vector pSM214 is used or vector pET-24b is used.
- 31. The method of claim 30, in which the protein of the invention is 953 and the vector is pSM214.
- **32.** A method for the heterologous expression of a protein of the invention, in which a protein is expressed or purified such that it adopts a particular multimeric form.
- 33. The method of claim 32, in which protein 953 is expressed and/or purified in monomeric form.
- 34. The method of claim 32, in which protein 961 is expressed and/or purified in tetrameric form.
- 35. The method of claim 32, in which protein 287 is expressed and/or purified in dimeric form.
- 15 **36.** The method of claim 32, in which protein 919 is expressed and/or purified in monomeric form.
  - **37.** A method for the heterologous expression of a protein of the invention, in which the protein is expressed as a lipidated protein.
- 38. The method of claim 37, in which the protein of the invention is 919, 287, ORF4, 406, 576, or ORF25.
  - **39.** A method for the heterologous expression of a protein of the invention, in which (a) the protein's C-terminus region is mutated and, optionally, (b) no fusion partner is used.
- 40. The method of claim 39, wherein the mutation is a substitution, an insertion, or a deletion
  - 41. The method of claim 40, wherein the protein of the invention is 730, ORF29 or ORF46.
  - 42. A method for the heterologous expression of a protein of the invention, in which the protein's leader peptide is mutated.
    - 43. The method of claim 42, in which the protein of the invention is 919.
    - 44. A method for the heterologous expression of a protein, in which a poly-glycine stretch within the protein is mutated.
- 45. The method of claim 44, wherein the protein is a protein of the invention.
  - 46. The method of claim 45, wherein the protein of the invention is 287, 741, 983 or Tbp2.
  - **47.** The method of claim 46, wherein (Gly)<sub>6</sub> is deleted from 287 or 983.
  - **48.** The method of claim 46, wherein (Gly)<sub>4</sub> is deleted from Tbp2 or 741
  - 49. The method of claim 47 or claim 48, wherein the leader peptide is also deleted.
- 45 **50.** The method of any preceding claim, in which the heterologous expression is in an E.coli host.
  - **51.** A protein expressed by the method of any preceding claim.
  - 52. A heterologous protein comprising the N-terminal amino acid sequence MKKYLFSAA.

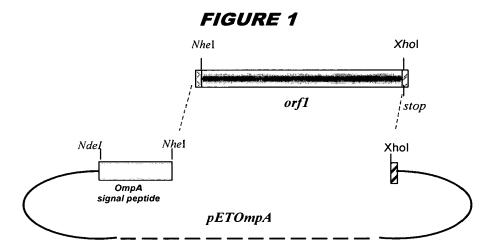
55

50

10

30

40



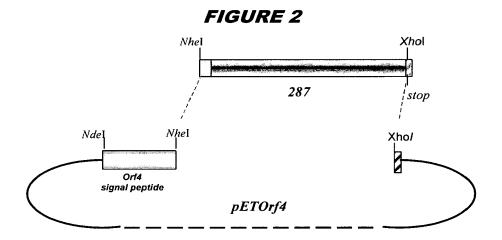
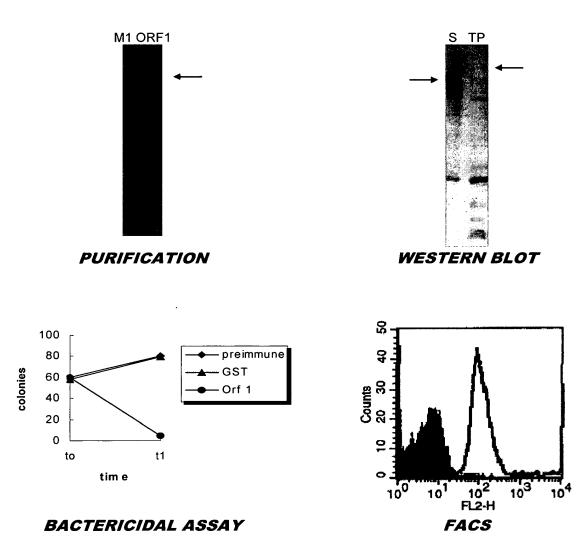
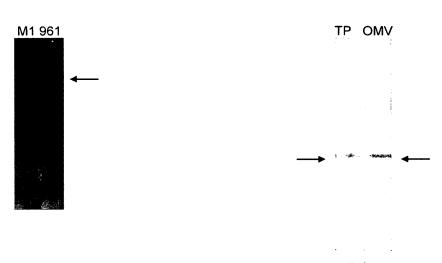


FIGURE 3

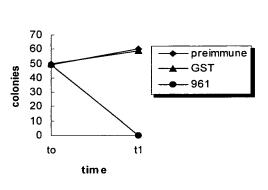


**ELISA:** POSITIVE

FIGURE 4



**PURIFICATION** 

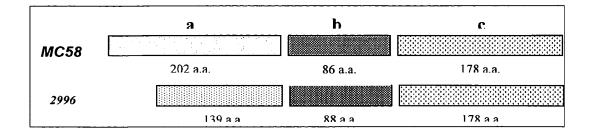


**BACTERICIDAL ASSAY** 

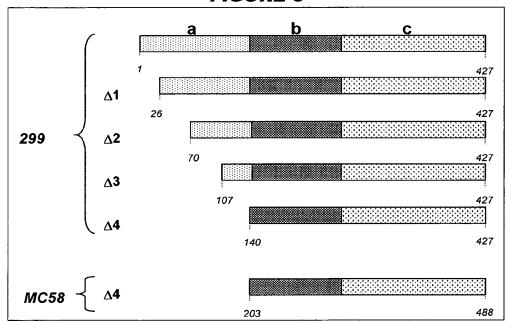
Oconties 100 101 102 103 104 FL2-H
FACS

**WESTERN BLOT** 

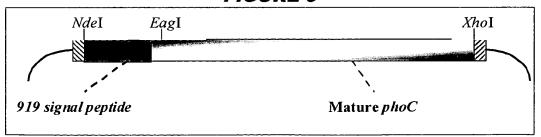
**ELISA:** POSITIVE

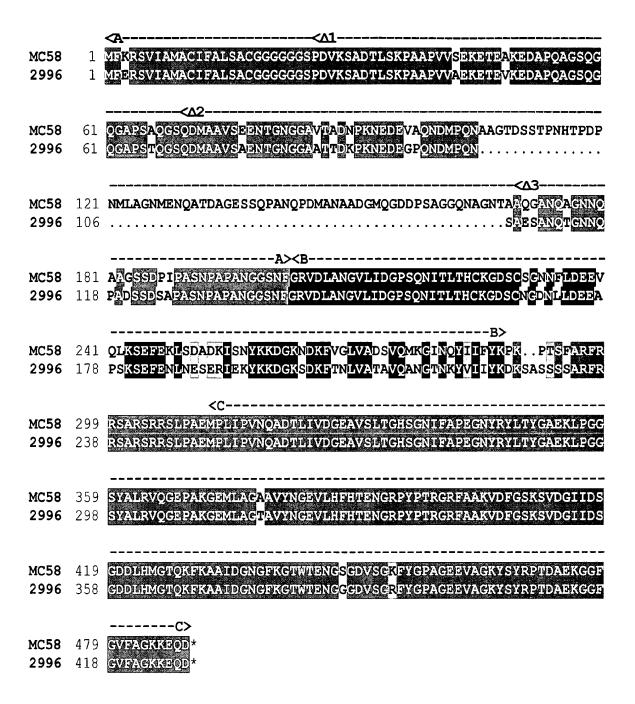


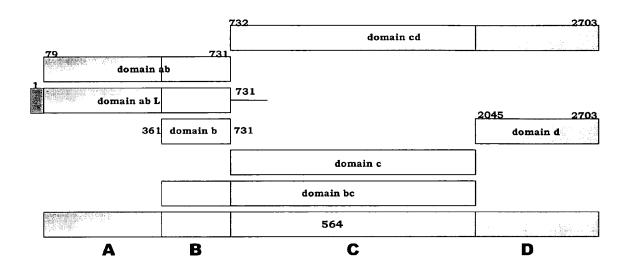
# FIGURE 6



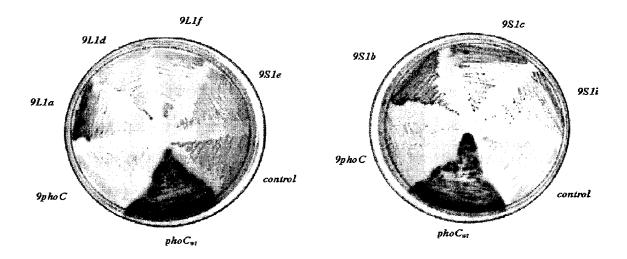
# FIGURE 9





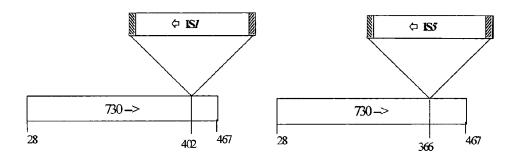


# FIGURE 10

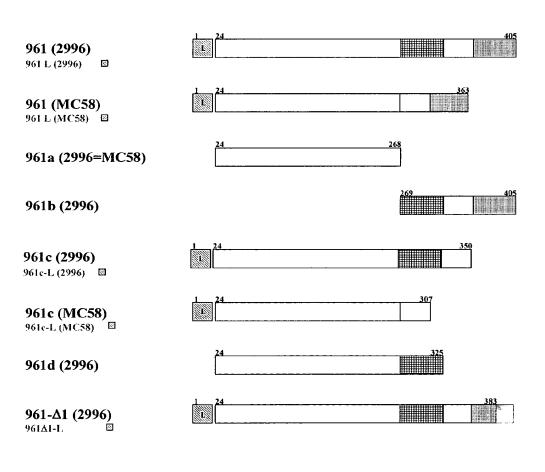


# FIGURE 11A

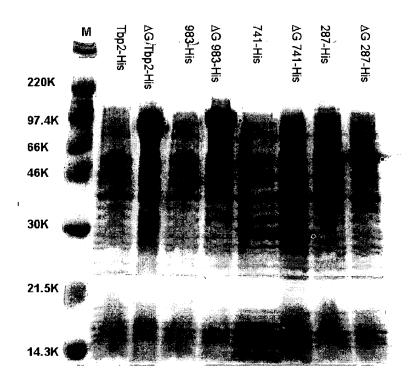
# FIGURE 11B



# FIGURE 12

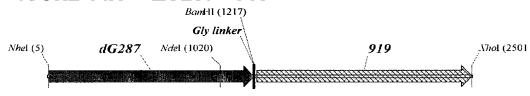


Membrane anchor

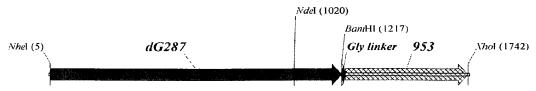


### FIGURE 14

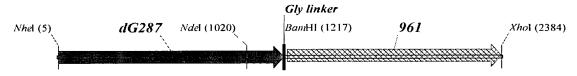
### FIGURE 14A — ΔG287—919



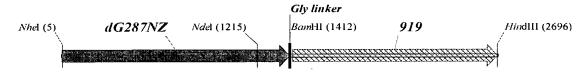
### FIGURE 14B — ΔG287—953



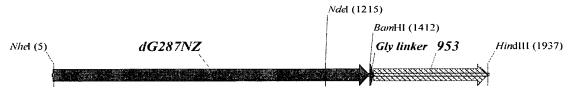
# FIGURE 14C - AG287-961



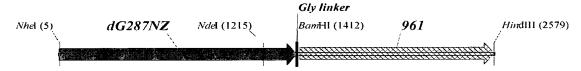
### FIGURE 14D — ΔG287NZ—919



# FIGURE 14E — ΔG287NZ—953



### FIGURE 14F — ΔG287NZ—961



### FIGURE 14G — ΔG983-ORF46.1



### FIGURE 14H — ΔG983-741



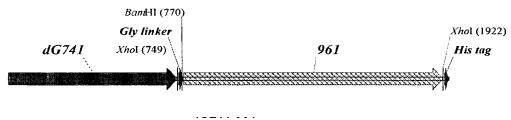
### FIGURE 14I — ΔG983-961



### FIGURE 14J — ΔG983-961c



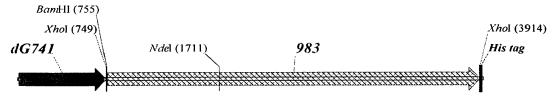
### FIGURE 14K — ΔG741-961



### FIGURE 14L — ΔG741-961c



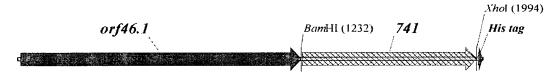
# FIGURE 14M — ΔG741-983



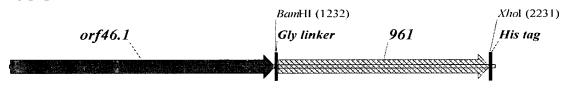
### FIGURE 14N — ΔG741-ORF46.1



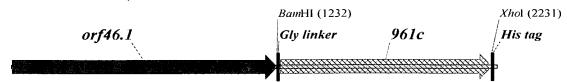
### FIGURE 140 — ORF46.1-741



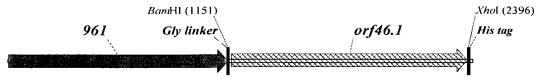
#### FIGURE 14P — ORF46.1-961



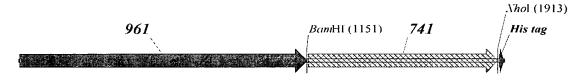
# FIGURE 14Q — ORF46.1—961c



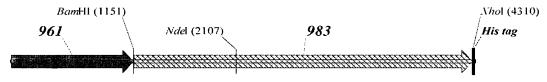
### FIGURE 14R — 961-ORF46.1



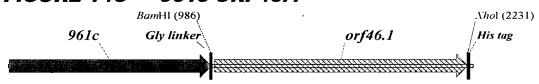
# FIGURE 14S — 961-741



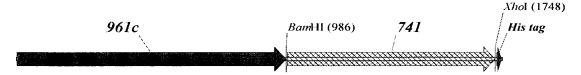
### FIGURE 14T — 961-983



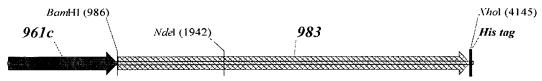
### FIGURE 14U --- 961c-ORF46.1



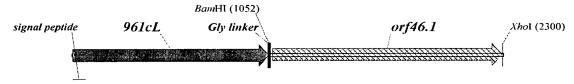
### FIGURE 14V — 961c-741



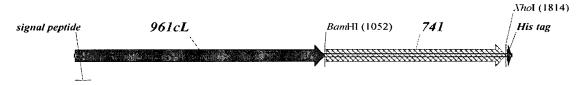
# FIGURE 14W — 961c-983



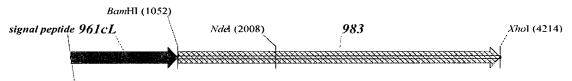
# FIGURE 14X — 961cL-ORF46.1



# FIGURE 14Y — 961cL-741



# FIGURE 14Z — 961cL-983



#### REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

#### Patent documents cited in the description

- WO 9924578 A [0002] [0004] [0004] [0005] [0005] [0149] [0178]
- WO 9936544 A [0002] [0004] [0004] [0005] [0005]
   [0178]
- WO 9957280 A [0002] [0004] [0004] [0005] [0005] [0102] [0167] [0173] [0178]
- WO 0022430 A [0002]

- WO 0066791 A [0006]
- WO 0066741 A [0008] [0035] [0037] [0053] [0074] [0086] [0101] [0130] [0158] [0172] [0194] [0279]
- WO 0071574 A [0009]
- WO 0104316 A [0009]
- WO 9955873 A [0148]

#### Non-patent literature cited in the description

- TETTELIN et al. Science, 2000, vol. 287, 1809-1815 [0006]
- URSINUS; HOLTJE. J.Bact., 1994, vol. 176, 338-343 [0111]
- HARZ. Anal. Biochem., 1990, vol. 190, 120-128 [0114]
- GLAUNER. Anal. Biochem., 1988, vol. 172, 451-464 [0117]
- THUNNISSEN et al. Biochemistry, 1995, vol. 34, 12729-12737 [0121]
- VAN ASSELT et al. Structure Fold Des, 1999, vol. 7, 1167-80 [0121]
- TERRAK et al. Mol.Microbiol., 1999, vol. 34, 350-64
   [0124]

- VELATI BELLINI et al. J. Biotechnol., 1991, vol. 18, 177-192 [0140]
- ROPP; NICHOLAS. J. Bact., 1997, vol. 179, 2783-2787 [0190]
- KRAFT. J. Bact., 1998, vol. 180, 3441-3447 [0191]
- THALLER et al. *Microbiology*, 1994, vol. 140, 1341-1350 [0248]
- HOICZYK et al. EMBO J, 2000, vol. 19, 5989-99
   [0272]
- EVANS. Infect.Immun., 1974, vol. 10, 1010-1017
   [0283]
- FILIP. J.Bact., 1973, vol. 115, 717-722 [0307]
- DAVIES. J.Immunol.Meth., 1990, vol. 143, 215-225
   [0307]